INFORMATION AND DECISION-MAKING IN SOCIO-BIOLOGICAL MULTI-AGENT SYSTEMS

A Dissertation
Presented to
The Academic Faculty

by

Keith C. Paarporn

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
May 2018

INFORMATION AND DECISION-MAKING IN SOCIO-BIOLOGICAL MULTI-AGENT SYSTEMS

Approved by:

Dr. Joshua S. Weitz, Advisor
Professor, School of Biological Sciences
Georgia Institute of Technology

Dr. Jeff S. Shamma
Professor, Computer, Electrical, and Mathematical Sciences and Engineering Division
King Abdullah University of Science and Technology

Dr. Yorai Wardi, Co-advisor
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. William C. Ratcliff
Assistant Professor, School of Biological Sciences
Georgia Institute of Technology

Dr. Magnus Egerstedt
Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Matthieu R. Bloch
Associate Professor, School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: March 29, 2018

ACKNOWLEDGEMENTS

There are four individuals I would like to thank that have had a profound academic influence on me during my time at Georgia Tech. First, I would like to thank my advisor, Joshua Weitz. Coming into the graduate program, I had no way of knowing I would end up doing an interdisciplinary PhD at the interface of biology and engineering. I had every intention to study control theory, joining Jeff Shamma's group during my first year. Joshua and Jeff were PI's on the ARO grant I was funded on, and when Jeff moved on to join the ranks of KAUST, I joined the ranks of the Weitz group. This new environment posed difficulties and challenges - why did I, a control theorist at heart, belong in a scientific research group? Joshua understood there were inherent differences between our worlds and gave me the space and creative freedom to succeed. Soon enough, Joshua and I began to speak the same language, making the PhD research experience all the more enjoyable. Joshua is a leader in his field of research and in training the new crop of scientists and engineers. He leads by example, being a model of efficiency and excellence, and constantly demonstrating how the quality and integrity of one's research can always be improved upon. He unfailingly supports his students and postdocs, both financially and as a dependable, trustworthy counselor for any problems that may have arisen.

I would like to thank my co-advisor Yorai Wardi. I was very fortunate that Yorai took an interest in my research, where he was able to provide a control-theoretic grounding in the scientific problems we were working on. I wish I had more time to soak up his worldly wisdom on function spaces, relaxed controls, and the finest foods across the globe.

I would like to thank Jeff Shamma, who introduced me to the academic world

and supervised my Master's thesis. Jeff is a renowned leader in his field, and an endless source of knowledge, insight, ideas, and humor. He possesses a long list of "half-baked" ideas, each of which could easily become the next student's PhD thesis. It would take several lifetimes to pick the vast depths of his brain. Although we interacted less frequently due to being half a world apart, I could still rely on Jeff for sound advice and feedback.

Lastly, I would like to thank Ceyhun Eksin. A close mentor, collaborator, and friend, Ceyhun provided the guidance I needed to navigate the interdisciplinary environment I found myself in. Our collaboration has been incredibly fruitful, and hopefully it continues into the future. I believe this is because no one else was willing to drop everything and spend hours at a time doing math on a whiteboard with me. Ceyhun is immensely supportive, constantly encouraging me to present and publish our work, and giving timely advice on all aspects of life. In the next stage of my career, I strive to honor the impact Ceyhun has had on my academic development by mentoring the next cohort of young researchers.

There are many, many more to thank. I will try my best to mention them all. The members of the Weitz group: Ashley Coenen, Devika Singh, Stephen Beckett, David Demory, Joey Leung, Yu-Hui Lin, Daniel Muratore, Shenyun Peng, Hend Alrasheed, Rong Jin, Adam Zhang, Brighton Ancelin, Walker Gussler, Bradford Taylor, Luis Jover, Alex Bucksh, Chad Wigington, and Hayriye Gulbudak - thanks for making the grad school experience so memorable. I want to thank members of Jeff Shamma's group: George Kotsalis, Kwang-Ki Kim, Behrouz Touri, Daniel Pickem, Farhan Aziz, and special thanks to Sebastian Ruf for his friendship, collaboration, and for making the dullest moments completely ridiculous. Outside of my immediate lab groups, I'd like to thank Sid Mayya, Maria Santos, Li Wang, and Gennaro Notomista as well as all the members of the GRITS lab. Additionally, I want to thank the members of my

committee - Magnus Egerstedt, Matthieu Bloch, and Will Ratcliff, for their invaluable feedback that helped me get to the finish line. I have also had the pleasure of collaborating with Philip Paré of UIUC on fascinating work regarding social dynamics. I would like to acknowledge the sources of funding that supported me during my graduate studies. A generous Army Research Office grant allowed me to work on the core of my thesis with intellectual curiosity. I have also had numerous opportunities as a teaching assistant in the school of ECE, which were great experiences for me.

Outside of the lab, I'd like to thank Marcus Daum, Jacqui Olson, and Conan Zhao for numerous Pho-related gatherings. I'd like to thank Paul Dorn, Michael Lu, and Ilya Perepelitsky for much-needed tennis breaks. I'd like to thank Philip Young and Pardee Poonjumnern for always being there. Special thanks to Jean Kuo and her family for their unwavering hospitality and kindness.

Above all, I'd like to thank my family. They support me without question. My mother is the most rational decision-maker I know. I benefited from her constant reminders of what is important, and that health is a priority second to none. My father's curiosity about my research stimulated discussions of what it means to be an engineer. I want to thank my brother James and sister-in-law Ashley for keeping me well-fed and being an important part of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The objective of this thesis is to deepen understanding of the role of information in multi-agent biological and social systems, and to identify ways that information can be influenced in order to produce a desirable outcome. These systems consist of a population of individual decision makers whose individual interests align with a common goal. Individuals may only have access to a limited amount or certain types of information and hence make only partially informed decisions. The global outcomes in these systems are therefore primarily determined by how information is distributed across the population.

We investigate the role information plays in three distinct multi-agent systems that appear in biological and social contexts. First, we study noisy communication in social dilemmas of cooperation. We analyze a game of incomplete information to identify the circumstances under which information sharing between players promotes coordination. Such social dilemmas are central to the study of evolutionary biology, microbial cooperation and competition, and collective animal behavior. Second, we study how individual awareness helps reduce the spread of a disease in a networked population. The individuals receive partial information about the epidemic through socially local as well as global sources, and take social distancing measures. We evaluate the dynamical behavior, and find desirable distributions of these information sources that mitigate the severity of an epidemic. Third, we consider how information and incentives can be influenced to promote the conservation of public resources. Self-interested individuals consume public goods with limited restraint, potentially leading to a tragedy of the commons. We formulate optimal control problems to allow an influential entity (e.g. the media, public figures) to perturb a population's perception

of the environmental state, with the objective of maximizing the accumulation of public goods over time. Using a suitable optimal control algorithm, we numerically characterize locally optimal policies to achieve this objective.

CHAPTER I

INTRODUCTION

A multi-agent system is comprised of an interacting population of individual, autonomous, and informed decision makers. Collectively, the individuals produce complex global behaviors. Three important examples of multi-agent systems in biological and social contexts are the dilemmas of cooperation, the spread of disease in a population, and the utilization of public resources. Central to the operation and dynamics of these three systems is the flow of information.

Information can be defined as any source or process that ellicits a functional reaction from its receiver [49]. It allows individuals to take strategic actions to protect their own interests. Information comes in many forms - knowledge, cues, signals, opinions, and combinations thereof. How information is distributed or exchanged among the agents determines their individual actions, ultimately leading to a global outcome. Therefore, information is a primary factor to understanding interactions and behaviors in complex systems.

The objectives of this thesis are to analyze how the distribution of information in the three aforementioned multi-agent systems drives decision-making and dynamics, as well as to identify ways to influence the distribution in order to steer the dynamics to a desirable outcome. By studying the role of information in these systems, this work furthers understanding in their respective application domains - evolutionary biology, epidemiology, and game theory. In particular, 1) we highlight the importance of information sharing as a factor for cooperation in social dilemmas, 2) the combination of epidemic information and social distancing alone is insufficient to eradicate an epidemic, and 3) the influence of public opinion can help conserve common resources.

Below, we provide the context and motivation for studying the role of information in these three settings.

1.1 The role of information in collective behavior

Groups of organisms across the animal kingdom exhibit cooperative behavior to survive and thrive in uncertain and fluctuating environments. They accomplish this despite selfish incentives to defect and despite imperfect informational capabilities [10, 23, 24]. A theoretical understanding of why collective behaviors emerge and are maintained starts at the individual level. How does the social context of the group interaction influence decision making at the individual level? What information does an individual have to base its decision on?

A notable example that highlights these features is quorum sensing in bacteria. Quorum sensing is a cell-to-cell communication system that operates by secreting and sensing signaling molecules called autoinducers [63, 41]. Quorum sensing allows bacteria to collectively respond to changes in their environment. Bacteria are highly social organisms, able to both sense the environment directly as well as to send and receive autoinducer signals. Once a cell detects a threshold concentration of autoinducer molecules is reached, indicating sufficiently high local cell density, it responds by activating gene expression. The benefit of group coordination relies on the participation of many cooperating cells. Cooperative behaviors such as the formation of biofilms, bioluminescence, motility, and secretion of virulence factors are facilitated by quorum sensing systems [63, 120, 67, 26].

How and why quorum sensing systems in bacteria have evolved remains largely unexplored [67, 86, 121]. Beyond the microbial world, decision-making in collective animal groups balance both private and socially obtained information [5, 64]. For example, honeybees and ant colonies alike aggregate privately obtained and social signals to collectively decide upon the best nesting site [32]. Swarming behaviors in

ants, fish, and birds are a result of decentralized communication and decision-making [22]. The underlying principles of observed collective decision-making processes is the subject of ongoing research [22, 100]. Motivated by these questions, our work rigorously addresses the relationship between social communication and individual decision-making strategies in group coordination social contexts. We accomplish this by analyzing a coordination game where each player is uncertain about the environmental state through sensed and shared signals, where we characterize the Nash equilibria and fitness maximizing strategies. Of note, we find one novel "majority logic" strategy that exploits social information to maximize cooperation - only in the scenario where players have access to both private and shared information.

The results presented for this work (Chapter 2) are currently under review for journal publication. A pre-print [74] can be found at http://dx.doi.org/10.1101/268268.

1.2 Mitigation of epidemic spreading through awareness

With the widespread availability of social media and news outlets on the internet and television, individuals may be well-informed about the current state of ongoing disease epidemics and how to take precautionary measures to avoid getting sick. In the recent 2009 H1N1 Influenza pandemic, people responded to public service announcements by increasing the frequency of washing hands, staying at home when they or loved ones were sick, or avoiding large public gatherings [105]. In the 2014 Ebola outbreak in West Africa, a combination of quarantining and sanitary burial methods were shown to significantly reduce the rate of virus spread [80]. These precautions and social distancing actions effectively limit epidemic spread. Individuals' distancing actions depend on the extent of how informed they are. The dissemination and exchange of information influences the public's behavior, affecting the course of the epidemic itself, which in turn affects the public's behavior again [35, 14]. This feedback loop allows epidemic spreading to coevolve with information and human social behavior,

inducing complex dynamics not observed by studying each process in isolation. A better understanding of what these complex dynamics are can help inform more effective implementations of public health campaigns and epidemic control strategies.

Models of epidemic spreading over networks have rapidly advanced in the last 15 years [114, 62], but human behavioral elements and the role of information in epidemics are only beginning to be studied in the same mathematical framework [33]. In this work, we formulate a susceptible-infected-susceptible (SIS) model by considering a networked population that receives dynamically distributed information about the epidemic from two sources. Individuals are partially informed about both the infection status of their local social contacts and of the population, and follow social distancing actions based on their personal information. We analyze the equilibrium and long-run properties of this system and compare these results to the benchmark models (no awareness). We find that awareness and social distancing reduces the endemic equilibrium state, but is insufficient to eradicate an epidemic. Through simulations, we find that local contact information is crucial in slowing the spread of an epidemic.

The results presented for this work (Chapter 3) are published [75, 76, 77].

1.3 Opinion influence for conservation of public goods

A "tragedy of the commons" [40] is a scenario where the selfish, myopic interests of individuals can eventually lead to the depletion of a shared common environment or public good. These situations are prevalent among social and biological systems across scales [20, 82, 90, 89]. The growing human population intensifies these concerns. By the year 2100, the United Nations estimates the world population will reach 11.2 billion[1] . This calls into question society's ability to support this projected number, given current consumption rates and detrimental effects on the environment due to

[1]http://www.worldometers.info/world-population/world-population-projections/

human activities [20]. Fisheries are experiencing reductions of 83% in population size [46] over the last 50 years, primarily due to overfishing.

To prevent such tragedies from occurring, researchers in the social sciences have called for interventions from a variety of instituions in society [73, 8, 113]. Information plays an important role in these interventions. Raising awareness through environmental activist movements as well as educating the public about current environmental conditions is necessary to effect change [82]. However, the efficacy of these proposed solutions are rarely tested using realistic quantitative models [73, 82]. In this thesis, we provide such an analysis by building upon a recent model of game-environment feedback dynamics [119] to include public opinions. We model the influence a central entity has on public opinion through information and incentive control policies that seek to conserve the environmental state. Using optimal control algorithms, we numerically characterize locally optimal policies that achieve this objective. We find that these policies encourage an "oscillating tragedy of the commons" [119], where the environment cycles between replete and deplete states. This resulting dynamic is undesirable in many cases, as it destabilizes the state of the commons.

The material presented for this work (Chapter 4) are both published [119] and currently under review for conference publication. A pre-print [78] can be found at https://arxiv.org/abs/1803.06737.

CHAPTER II

THE ROLE OF INFORMATION IN COLLECTIVE BEHAVIOR

In stochastically fluctuating environments, organisms modify their behavior to increase their chances of growth and survival. For instance, bacteria can sense the extracellular state of resources, toxins, and even the presence of other bacteria. Individual cells can then rapidly change their behavior, i.e. their "phenotype". Such stochastic phenotype switching may be random and unrelated to sensed cues - this strategy is known as bet-hedging. Or, such stochastic phenotype switching can be adaptive, and directly in response to sensed cues - this strategy is called phenotypic plasticity [57, 84]. However in many circumstances, the chance that one cell grows or survives depends critically on the behaviors of other cells. In those circumstances, coordination of behavior with others directly influences survival.

In bacteria, some cooperative behaviors are facilitated by quorum sensing, a cell-to-cell communication system in which individual cells secrete and sense autoinducer molecules to obtain information about the environment and to gauge local cell density [63, 41]. Quorum sensing can induce formation of biofilms for protection against a host's immune system, secretion of virulence factors to consolidate colonization of the host, motility, and many others [120, 67, 26, 9]. In honeybee and ant colonies, individuals process and share information to collectively reach an informed collective decision about the best nesting site [32, 87]. The individual-level mechanisms that produce collective behaviors in animal groups is an area of ongoing research [22, 100]. Inspired by these examples, the aim of this chapter is to develop a game-theoretic framework in which to study the individual-level decision-making processes which

produce collective behavior under environmental uncertainty and noisy communication.

There is an extensive body of work on individual (as opposed to group) decision-making in fluctuating environments [84, 57]. In many scenarios, an individual must match its phenotype or behavior to changing conditions by using sensory cues from the environment, signals from other individuals in the population, or both [54, 53, 93, 27, 28]. An individual employing the optimal bet-hedging strategy diversifies behaviors at frequencies that mirror the posterior knowledge of environmental fluctuations. These results quantify an information-theoretic connection between the optimal fitness value and the amount of information about the environment available to the individual. However, the resulting optimal individual fitness is independent of the actions of others in the population. Hence, these models do not address the interplay between information and collective decision-making in fluctuating environments. In bacterial quorum sensing, the autoinducer signaling molecules that individual cells send serve as social cues that indicate local cell density [63, 67, 86]. Substantial experimental research has been done in recent years to unravel the mechanisms of this complex communication system. Nonetheless, questions remain regarding why such communication systems are utilized, particularly from an evolutionary standpoint [121].

Game theory offers a framework to explain rational behaviors when an individual's well-being depends on the actions of others in the population. To understand the role of communication in collective decision-making under uncertainty, we recognize there are two components to the decision-making process: 1) a *communication system*, or the way individuals acquire information and 2) *strategies*, or the way individuals use acquired information to make a decision. In this vein, two recent works have attempted to understand the role of communication in group coordination from a game-theoretic perspective. In both studies, the players' strategies are assumed to

be fixed while the communication systems are evolvable. [79] studies the evolutionary outcomes of communication systems in an N-person volunteer's dilemma game, where all players adopt a majority rules strategy. [19] consider how the choice of communication systems affect levels of cooperation between two populations of microbes in an information exchange game. In that model, each individual follows a bet-hedging strategy.

Complementary to these studies, in this chapter we seek to find the optimal strategies that promote coordination given the limitations of a fixed, noisy communication system. We study a two-player, two-action game in which the environment stochastically switches between two possible states. The state determines the game payoffs, and players must coordinate on the correct action corresponding to the environment. We study two versions of this game in order to highlight the value of information sharing. In the first, players only receive a private cue from the environment. In the second, players receive a shared social cue in addition to their private cue. In doing so, we find three optimal strategies that emerge given variation in information fidelity. Most notable is a "majority logic" type strategy that is optimal when environmental sensing has intermediate reliability, and information sharing is very reliable. As we show, this allows an individual to act upon their inference of the environmental state when it is validated by social information.

2.1 A coordination game in fluctuating environments

We consider a two action, two player game that is played repeatedly in stages $t = 1, 2, \ldots$. We denote the set of players $\mathcal{N} \equiv \{1, 2\}$ with generic member i, and the set of actions $\mathcal{X} \equiv \{A, B\}$. At each stage t, the environment $E(t)$ takes one of two possible states, e_A or e_B. If $E(t) = e_A$ then $E(t+1) = e_B$ with probability v_{AB}, and if $E(t) = e_B$ then $E(t+1) = e_A$ with probability v_{BA}. We assume that $E(1)$ is arbitrarily determined and the switching probabilities $v_{AB} > 0$ and $v_{BA} > 0$ are

fixed. Hence, the environmental state evolves according to a two-state Markov chain that spends a fraction $v_A \equiv \frac{v_{BA}}{v_{BA}+v_{AB}}$ and $v_B \equiv \frac{v_{AB}}{v_{BA}+v_{AB}}$ of time in states e_A and e_B, respectively. Once $E(t)$ is determined, players select their actions $x_i(t) \in \mathcal{X}$. They do not know $E(t)$ with certainty, and the state $E(t)$ governs the players' utility functions at each stage. We are interested in the average long-term fitness of the players given the stochastically switching environment.

Before defining long-term fitness, we define the game that is played at each stage as follows. Both must coordinate on action A (B) if $E(t) = e_A$ (e_B) to receive a payoff $b_A > 0$ ($b_B > 0$). If they miscoordinate, or if they coordinate on the incorrect action, the payoff is zero. The environment-dependent payoff matrices are illustrated in Figure 2.1. Note that the payoff value to either player is the same. Thus, a generic player's utility function is defined as

$$U(x_1, x_2, E(t)) \equiv \begin{cases} b_A & \text{if } x_1 = x_2 = A \text{ and } E(t) = e_A \\ b_B & \text{if } x_1 = x_2 = B \text{ and } E(t) = e_B \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

Given the realization of the environment $E(t)$, a normal form game is played at stage t between the set of players $\mathcal{N}$ equipped with action set $\mathcal{X}$, and with utility function $U(\cdot, \cdot, E(t))$. We denote this by the triple $\mathcal{G} \equiv (\mathcal{N}, \mathcal{X}, U(\cdot, \cdot, E(t)))$ (this follows standard game-theoretic notation to denote normal-form games). In the following two subsections, we introduce two models where players receive noisy information about $E(t)$, and present our definition of long-term average fitness.

2.1.1 The game $\mathcal{G}_p$ with environmental sensing

At each stage t, suppose each player ($i = 1, 2$) independently senses the environment $E(t)$ by receiving a private cue $\alpha_i(t) \in \{e_A, e_B\}$. The cue $\alpha_i(t)$ matches the true environmental state with probability $p \in [1/2, 1]$, and mismatches with probability

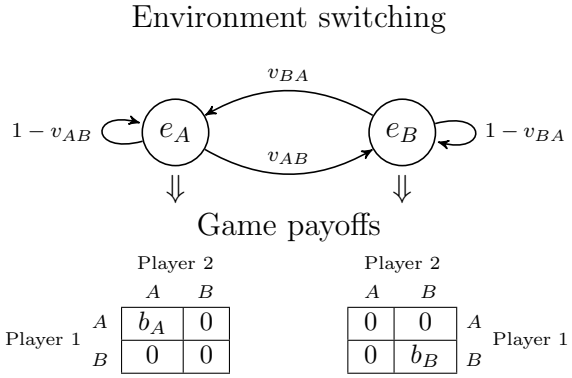


Figure 2.1: After each stage, the environment stochastically switches between e_A and e_B and determines the payoff matrix. In e_A (e_B), players need to coordinate on the A (B) action to receive a positive payoff b_A (b_B).

$1 - p$.

$$P(\alpha_i(t) = X | E(t) = Y) = \begin{cases} p & \text{if } X = Y \\ 1 - p & \text{if } X \neq Y \end{cases}, \; X, Y \in \{e_A, e_B\}$$

Thus, $\alpha_i(t)$ is the output of a binary noisy channel of fidelity p, NC(p), whose input is $E(t)$. We will refer to the parameter p as the sensing fidelity.

A *strategy* is a mapping from the set of cues $\{e_A, e_B\}$ to the set of actions $\mathcal{X}$. In other words, a strategy is a contingency plan or policy that a player adopts that assigns the action to take upon receiving a particular cue. Hence, it is a description of how a player makes informed decisions. Since there are two possible cues and two actions, each player can choose among $2^2 = 4$ strategies. We assume player i chooses only one strategy s_i, which is fixed for all stages $t = 1, 2, \ldots$. For notational convenience, we will denote a strategy by an ordered two-vector whose entries are either A or B. For instance, $s_i = [A, B]$ denotes the strategy where i plays action A when $\alpha_i(t) = e_A$ and action B when $\alpha_i(t) = e_B$. Additionally, we write $s_i(e_A) = A$ and $s_i(e_B) = B$. We denote the set of all four strategies $\mathcal{S}^4$.

The list of all four strategies is given in Table 1. They are read "Only A" (OA), "Follow Cue" (FC), "FC bar" ($\overline{\text{FC}}$), and "Only B" (OB).

Remark 2.1. *These strategies, with the exception of $\overline{FC}$, can be classified as a max*

Table 1: List of all strategies in the game with no information sharing.

$\alpha_i/\ s_i$	s_{OA}	s_{FC}	$s_{\overline{\mathrm{FC}}}$	s_{OB}
e_A	A	A	B	B
e_B	A	B	A	B

likelihood estimate of the environmental state, depending on the environmental switching probabilities v_A, v_B and private cue fidelity p. Specifically, the strategy

$$s_{MLE}(\alpha_i) = \arg\max_{X=A,B} P(E = e_X|\alpha_i)$$

is precisely OA ($v_B < p < v_A$), OB ($v_A < p < v_B$), or FC ($v_A < p, v_B < p$). In this chapter, we omit analysis of the s_{MLE} strategy because it simply corresponds to one strategy in Table 1 for any given set of parameters.

When $p = 1/2$, s_{FC} and $s_{\overline{\mathrm{FC}}}$ reduce to the strategy that uniformly randomizes between A and B.

A measure of long-term fitness is the fraction of time the players coordinate on the correct action, weighted by the benefits b_A and b_B accordingly. We calculate this measure as follows. Consider the environment $E(t)$, which we recall evolves according to the two-state Markov chain with stationary distribution (v_A, v_B). Additionally, the players' cues are drawn independently of each other, but conditionally on the state $E(t)$. Since neither the cues nor player actions affect the environment switching probabilities, the fraction of time spent in the aggregate state (α_1, α_2, E) is given by the following stationary distribution

$$\pi_p(\alpha_1, \alpha_2, E) = \begin{cases} v_A P(\alpha_1|e_A)P(\alpha_2|e_A), \text{ if } E = e_A \\ v_B P(\alpha_1|e_B)P(\alpha_2|e_B), \text{ if } E = e_B \end{cases} . \tag{2.2}$$

For instance, the value of the stationary distribution π_p at the entry (e_A, e_A, e_B) is $v_A p(1-p)$. We define the *average long-term fitness* as the expected utility with

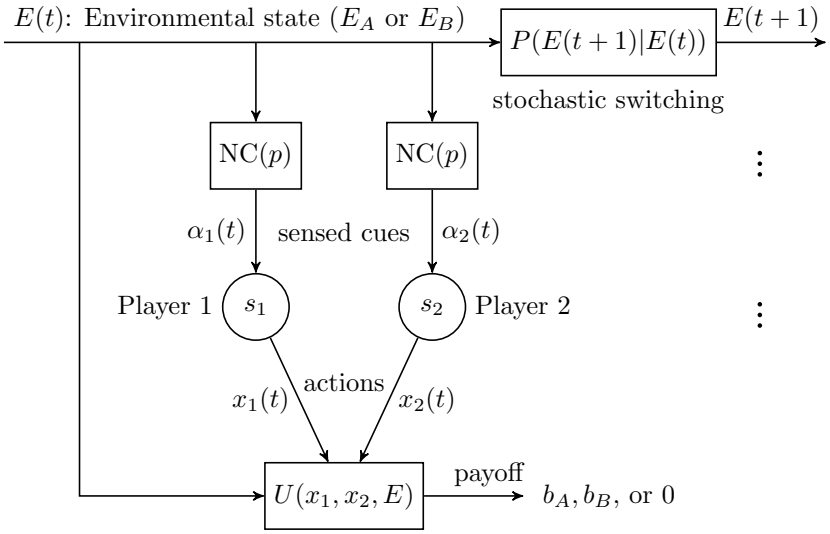


Figure 2.2: Diagram of the game with no information sharing $\mathcal{G}_p$. At stage t, each player independently senses a cue $\alpha_i(t)$ directly from the environment $E(t)$ through a binary noisy channel with fidelity p, NC(p). That is, $\alpha_i(t) = E(t)$ with probability p, and $\alpha_i(t) \neq E(t)$ with probability $1 - p$. Player i's strategy $s_i(\alpha_i(t))$ determines its action $x_i(t) \in \{A, B\}$. The payoff $U(x_1, x_2, E)$ to both players is either b_A, b_B, or 0, which is determined by the current actions of both players and the current environment. The environmental state stochastically switches to $E(t+1)$ at stage $t+1$.

respect to the stationary distribution,

$$f_p(s_1, s_2) \equiv \sum_{\alpha_1, \alpha_2, E} \pi_p(\alpha_1, \alpha_2, E) U(s_1(\alpha_i), s_2(\alpha_j), E). \tag{2.3}$$

Via (2.3), we define the normal form game $\mathcal{G}_p \equiv (\mathcal{N}, \mathcal{S}^4, f_p)$ played between players $\mathcal{N}$ equipped with strategy space $\mathcal{S}^4$, and with utility function f_p for a given sensing fidelity $p \in [1/2, 1]$. Figure 2.2 shows a diagram of the communication system underlying $\mathcal{G}_p$. We can represent $\mathcal{G}_p$ with the following 4×4 payoff matrix of long-term fitnesses $f_p(s_1, s_2)$,

$$\begin{array}{c} \\ s_{\mathrm{OA}} \\ s_{\mathrm{MLE}} \\ s_{\overline{\mathrm{MLE}}} \\ s_{\mathrm{OB}} \end{array} \begin{array}{c} \begin{array}{cccc} s_{\mathrm{OA}} & s_{\mathrm{MLE}} & s_{\overline{\mathrm{MLE}}} & s_{\mathrm{OB}} \end{array} \\ \begin{bmatrix} c_A & c_A p & c_A \bar{p} & 0 \\ c_A p & (c_A + c_B) p^2 & (c_A + c_B) p\bar{p} & c_B p \\ c_A \bar{p} & (c_A + c_B) p\bar{p} & (c_A + c_B) \bar{p}^2 & c_B \bar{p} \\ 0 & c_B p & c_B \bar{p} & c_B \end{bmatrix} \end{array} \tag{2.4}$$

where $\bar{p} \equiv 1 - p$ and

$$c_A \equiv b_A v_A \text{ and } c_B \equiv b_B v_B$$

are the relative benefits of each environmental state. An expression to calculate each entry of (2.4) is given in the Appendix 2.A. Note the payoff matrix is symmetric, which gives 10 unique strategy profiles. Therefore the identity of the players do not matter.

Remark 2.2. *The average long-term fitness* (2.3) *can also be viewed as the* ex-ante *expected utility in the one-shot Bayesian game $\mathcal{B}_p$ consisting of the set of players $\mathcal{N}$, action set A, external states $\{e_A, e_B\}$, type space $\{e_A, e_B\}$, utility U (eq.* (2.1)*), beliefs $P(\alpha_1, \alpha_2, E)$, and a common prior on E, $P(e_A) = v_A$ and $P(e_B) = v_B$. The ex-ante expected utility is defined as the expectation of U with respect to the belief $P(\alpha_1, \alpha_2, E)$, which is the induced probability distribution over the aggregate state of the world and coincides with* (2.2)*. In an ex-ante setup, players evaluate their*

utilities before receiving their signal α_i, and therefore must reason about the possible environmental states, the signal of the other player, and its own signal. The (normal form) Nash equilibrium solution concept applied to the game $\mathcal{G}_p$ coincides with the definition of ex-ante Bayesian Nash equilibrium (BNE) of $\mathcal{B}_p$. A Bayesian Nash equilibrium describes a state of rationality where no player can profitably deviate by changing its strategy given its belief of the world. We focus our attention on the normal-form formulation $\mathcal{G}_p$ in this chapter, keeping in mind that a Nash equilibrium in $\mathcal{G}_p$ can also be interpreted as a Bayesian Nash equilibrium of $\mathcal{B}_p$. A treatment of Bayesian games can be found in Ch. 6 of [111].

2.1.2 The game $\mathcal{G}_{pq}$ with information sharing

The game $\mathcal{G}_p$ is extended by allowing players to share their private cues with each other before deciding on an action. At stage t, after players sense $\alpha_i(t)$, player i sends a social cue $\beta_j(t)$ to player j ($j \neq i$), which matches i's private cue $\alpha_i(t)$ with probability $q \in [1/2, 1]$.

$$P(\beta_i(t) = X | \alpha_j(t) = Y) = \begin{cases} q & \text{if } X = Y \\ 1 - q & \text{if } X \neq Y \end{cases}, \; X, Y \in \{e_A, e_B\}$$

Thus, $\beta_i(t)$ is the output of a binary noisy channel of fidelity q, NC(q), whose input is $\alpha_j(t)$. We will refer to parameter q as the *sharing fidelity*. We assume here that players signal honestly, i.e. each player attempts to share their true private cue. This is a reasonable assumption because the players' interests are aligned - they must attempt to coordinate. In different contexts where player interests conflict, dishonest signalling becomes a rational alternative. For example, male fiddler crabs with inferior claws can bluff fighting ability to ward off other males [11].

Player i's information is now composed of the pair $y_i(t) \equiv (\alpha_i(t), \beta_i(t)) \in \{e_A, e_B\}^2$. Similarly for this model, a strategy is a mapping from the set of information pairs

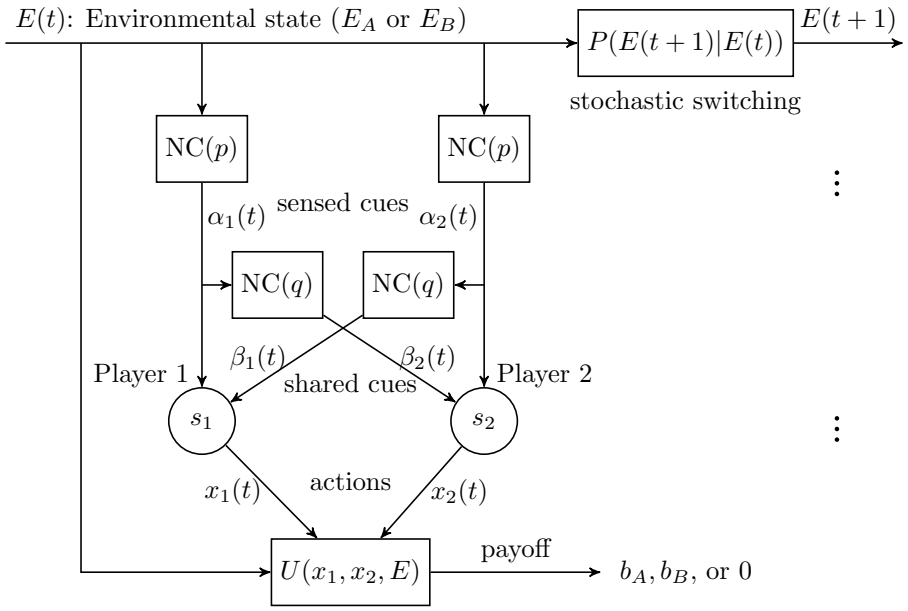


Figure 2.3: Diagram of the game with information sharing $\mathcal{G}_{pq}$. At stage t, each player independently senses a private cue $\alpha_i(t)$ directly from the environment $E(t)$ through a binary noisy channel with fidelity p, NC(p). Then, each player receives a signal $\beta_i(t)$ from the other's private cue through a separate noisy channel of fidelity q, NC(q). Given player i's information $(\alpha_i(t), \beta_i(t))$, its action is determined by its strategy, $s_i(\alpha_i(t), \beta_i(t)) = x_i(t)$. The payoff $U(x_1, x_2, E)$ to both players is either b_A, b_B, or 0, which is determined by the actions of both players and the current environment. The environmental state stochastically switches to $E(t+1)$ at stage $t+1$.

$\{e_A, e_B\}^2$ to actions, so i can now choose among $4^2 = 16$ strategies. We represent strategies as ordered four-vectors whose entries are A or B. For instance, $s_i = [A, B, B, A]$ is the strategy where i plays action $a_i = A$ when $y_i = (e_A, e_A)$, B when $y_i = (e_A, e_B)$, B when $y_i = (e_B, e_A)$, and A when $y_i = (e_B, e_B)$. We note that the four strategies available in $\mathcal{G}_p$ are also strategies in this game. They are now represented by the vectors $s_{\text{OA}} = [A, A, A, A]$, $s_{\text{FC}} = [A, A, B, B]$, $s_{\overline{\text{FC}}} = [B, B, A, A]$, $s_{\text{OB}} = [B, B, B, B]$. These four strategies base decision-making either on no information at all (OA and OB), or only on the private cue α_i (FC and $\overline{\text{FC}}$). The 12 new strategies base decisions on both private and shared signals. The set of all 16 strategies is denoted $\mathcal{S}^{16}$.

Similarly to (2.2), we calculate the average long-term fitness by considering the aggregate state (y_1, y_2, E), which occurs a fraction of the time according to the following stationary distribution

$$\pi_{pq}(y_1, y_2, E) = \begin{cases} v_A P(y_1, y_2|e_A), \text{ if } E = e_A \\ v_B P(y_1, y_2|e_B), \text{ if } E = e_B \end{cases}.$$

One can expand $P(y_1, y_2|E) = P(\alpha_1|E)P(\alpha_2|E)P(\beta_1|\alpha_2)P(\beta_2|\alpha_1)$. For instance, the value of the stationary distribution at the entry $((e_A, e_A), (e_A, e_B), e_A)$ is $v_A p^2 q(1-q)$. The *average long-term fitness* at $(p, q) \in [1/2, 1]^2$ is defined as

$$f_{pq}(s_1, s_2) \equiv \sum_{y_1, y_2, E} \pi_{pq}(y_1, y_2, E) U(s_1(y_1), s_2(y_2), E). \tag{2.5}$$

Via (2.5) and following the arguments from Remark 2.2, we define the normal form game $\mathcal{G}_{pq} \equiv (\mathcal{N}, \mathcal{S}^{16}, f_{pq})$ played between the set of players $\mathcal{N}$ equipped with strategy space $\mathcal{S}^{16}$, and with utility function f_{pq} for a given pair of sensing and sharing fidelities $(p, q) \in [1/2, 1]^2$. Figure 2.3 shows a diagram of the communication system that underlies $\mathcal{G}_{pq}$. An expression to calculate each entry of the resulting 16×16 symmetric payoff matrix is given in Appendix 2.A. Given the symmetry of the payoff matrix, there are 136 unique strategy profiles. A summary of all relevant parameters that define the games $\mathcal{G}_p$ and $\mathcal{G}_{pq}$ is listed in Table 2.

Remark 2.3. *We note that the payoff matrices of $\mathcal{G}_p$ and $\mathcal{G}_{pq}$ are symmetric. Hence, the identity of the player (player 1 or 2) does not matter. Such games fall into the class of potential games [66, 98]. A game is a potential game if the incentives of the players align with a global* potential *function. Here, the change in payoff from a unilateral deviation of a single player is equivalent to the change in global potential, given all other players remain the same. In our formulation, the potential function is the average player fitness, or welfare.*

Table 2: List of parameters that define the games $\mathcal{G}_p$ and $\mathcal{G}_{pq}$.

Parameter	Description	Notes
e_{X}	environmental state, $\mathrm{X} = \{A, B\}$	
p	environment sensing fidelity	$\in [1/2, 1]$
q	sharing fidelity	$\in [1/2, 1]$
α_i	private environment cue for player i	$\in \{e_A, e_B\}$, $i = 1, 2$
β_i	social cue player i receives from player j	$\in \{e_A, e_B\}$, $j \neq i$
b_{X}	payoff of coordinating in e_{X}, $\mathrm{X} = \{A, B\}$	> 0
v_{AB}, v_{BA}	prob. of switching to e_B from e_A, vice versa	both $\in [1/2, 1]$
v_{X}	fraction of time spent in e_{X}, $\mathrm{X} = \{A, B\}$	$v_A + v_B = 1$
c_{X}	relative payoff of coordinating in e_{X}	$= v_{\mathrm{X}} b_{\mathrm{X}}$
κ	ratio of relative payoffs	$= c_A / c_B$

2.2 Nash equilibria and fitness maximizers

We consider the Nash equilibrium solution concept and fitness maximizing strategies as notions of rational and optimal behavior, respectively. The comparison of the fitness maximizing strategies between the two models ($\mathcal{G}_p$ and $\mathcal{G}_{pq}$) presented offers a principled way to investigate the role of sensing and communication in collective decision-making in systems such as quorum sensing.

Specifically, we are interested in how the quality of the communication system, defined by the parameter p for $\mathcal{G}_p$ and (p, q) for $\mathcal{G}_{pq}$, dictates which strategy profiles are Nash equilibria and fitness maximizers. A Nash equilibrium is the classical solution concept in game theory which describes a rational steady-state configuration where no player has an incentive to deviate from its strategy. A Nash equilibrium in $\mathcal{G}_p$ is a strategy profile (s_1^*, s_2^*) that satisfies

$$\begin{aligned} f_p(s_1^*, s_2^*) &\geq f_p(s_1, s_2^*) \\ f_p(s_1^*, s_2^*) &\geq f_p(s_1^*, s_2) \end{aligned} \tag{2.6}$$

for all $s_1 \in \mathcal{S}^4$, $s_1 \neq s_1^*$, and for all $s_2 \in \mathcal{S}^4$, $s_2 \neq s_2^*$. A strict Nash equilibrium is a Nash equilibrium satisfying (2.6) with strict inequality. We say a strategy profile

$(\hat{s}_1, \hat{s}_2)$ is a *fitness maximizer* at p if it satisfies

$$f_p(\hat{s}_1, \hat{s}_2) = \max_{s_1, s_2 \in \mathcal{S}^4} f_p(s_1, s_2).$$

The same definitions above apply for $\mathcal{G}_{pq}$, where f_p is replaced by f_{pq}, and $\mathcal{S}^4$ by $\mathcal{S}^{16}$. In our coordination game formulation, due to the fact that f_p or f_{pq} is the fitness value to both players, the following fact holds: In $\mathcal{G}_p$ and $\mathcal{G}_{pq}$, the fitness maximizer(s) is necessarily a Nash equilibrium.

However, the converse is not true. We note that in our model, there is no fitness cost to having higher sensing and sharing fidelities p and q. Our aim is not to investigate such evolutionary tradeoffs, but to identify the types of strategies that ensure coordination in fluctuating environments given that the players utilize a communication system of quality $(p, q) \in [1/2, 1]^2$.

2.2.1 Nash equilibria and fitness maximizers in $\mathcal{G}^p$

The base game $\mathcal{G}$ (see Eq. (2.1)) admits a single strict Nash equilibrium at the correct coordinated action. We might expect the game $\mathcal{G}_p$, represented by (2.4), also has a coordination structure. That is, the players are better off if they coordinate on the same strategy in $\mathcal{S}^4$. In other words, they should play the same action given they receive the same signal. We can show that this intuition is indeed correct and state the following result.

Proposition 2.1. *For $p \in [1/2, 1]$, all Nash equilibria and fitness maximizers of $\mathcal{G}_p$ are necessarily symmetric strategy profiles, i.e. $s_1^* = s_2^*$.*

The proof, which we show in Appendix 2.B, of this result relies on exhaustively showing that for any asymmetric strategy profile, one player can switch to the other player's strategy to improve the fitness f_p. This means that any entry (i, j) below the diagonal in (2.4) is exceeded by (i, i), (i, j), or both. Due to this result, we can limit our analysis to the four symmetric strategy profiles. To refer to symmetric strategy

profiles (s, s) for the four strategies we simply write "OA" to denote (s_{OA},s_{OA}), and similarly for the other three.

We find that OA and OB are always Nash equilibria regardless of the value for sensing fidelity p and the relative benefits c_A, c_B. We find that $\overline{\text{FC}}$ is not a Nash equilibrium for any set of parameters. Figure 2.4 (left) illustrates the Nash equilibrium region of FC. We can characterize the equilibrium region of FC in the game $\mathcal{G}_p$ with respect to just two parameters, $(p, \kappa) \in [1/2, 1] \times (0, \infty)$, where

$$\kappa \equiv \frac{c_A}{c_B}.$$

is the ratio of relative benefits c_A to c_B. This is possible by normalizing the fitness, f_p/c_B, and applying (2.6) to solve for the region's conditions on p and κ.

Figure 2.4 (right) illustrates the regions where OA, OB, and FC are fitness maximizers. While OA and OB are Nash equilibria everywhere, there is a unique fitness maximizer (OA,OB, or FC), except for on the boundaries dividing each region, for any given (p, κ) value in $[1/2, 1] \times (0, \infty)$. On the boundary lines, the fitnesses of the two strategy profiles that are separated coincide. We note here that the region where FC maximizes fitness is a subset of its Nash equilibrium region. This reveals some points in the parameter space where FC is a suboptimal Nash equilibrium to the OA/OB strategies.

The fitness for FC increases quadratically in p: $f_p(\text{FC}) = (c_A + c_B)p^2$. Hence, with higher sensing fidelity, players are better able to infer the correct environmental state and select the correct action. The p^2 term appears because each player is independently sensing E, and coordination depends on both players independently receiving the correct cue which occurs with probability p^2. The value of employing FC diminishes as $\kappa \to \infty$ or 0. In either extreme, one environment is favored over the other. Either one occurs far more frequently over the long run, or its coordination benefit outweighs that of the other state. Hence, players do best by adopting either OA or OB instead of FC in these situations. FC is most desirable when $\kappa \approx 1$, where

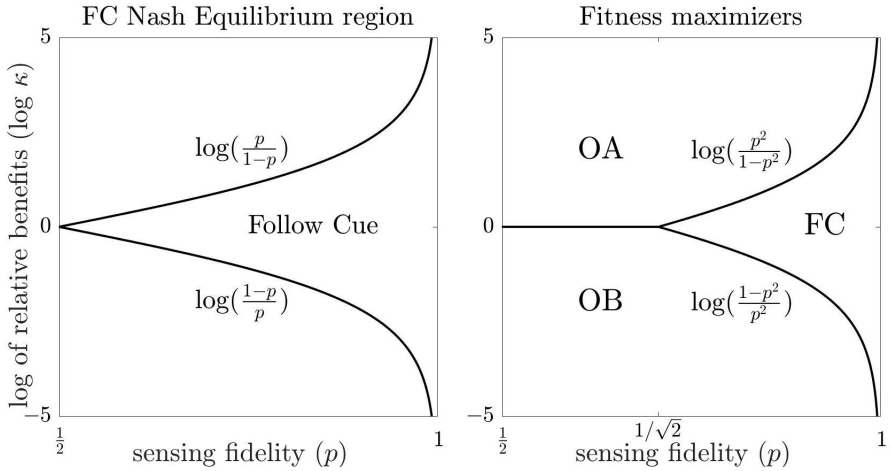


Figure 2.4: Characterization of strategies in the no information sharing game with respect to the parameter space (p, κ) with $\kappa = \frac{c_A}{c_B}$ ranging in log scale from -5 to 5. Log scale is used to show symmetry between the ranges $0 < \kappa < 1$ and $\kappa < \infty$. The quantity κ is the ratio of relative benefits between environment e_A and e_B. (Left) The region where FC is a strict Nash equilibrium : $\frac{1-p}{p} < \kappa < \frac{p}{1-p}$. The strategies OA and OB are Nash equilibria everywhere. (Right) Disjoint regions where the strategies OA, OB, and FC are fitness maximizers. OA: $\kappa > \max\left\{1, \frac{p^2}{1-p^2}\right\}$. OB: $\kappa < \min\left\{1, \frac{1-p^2}{p^2}\right\}$. FC: $\frac{1-p^2}{p^2} < \kappa < \frac{p^2}{1-p^2}$

either the environment fluctuates frequently and $b_A \approx b_B$, or a rare environment offers an enormous coordination benefit compared to the other. These are situations where acting on knowledge of the environment is most crucial. By playing OA or OB, players miss out on half of the fitness benefit opportunity whereas players employing FC are able to adapt to changing conditions.

Variants of s_{FC} are often identified as the optimal strategy that maximizes long-term fitness in fluctuating environments when individual fitness is uncoupled to the actions of others. They are often referred to as responsive switching [53], adaptive plasticity [29], and proportional betting [27]. In our model, the social context of coordination renders it suboptimal when private information is unreliable (low p), though it is still optimal for sufficiently high values of p.

2.2.2 Nash equilibria and fitness maximizers in $\mathcal{G}_{pq}$

We now turn to the game $\mathcal{G}_{pq} = (\mathcal{N}, \mathcal{S}^{16}, f_{pq})$. The Nash condition remains the same as in (2.6), except the long-term fitness is now given by f_{pq} (2.5). In comparison to $\mathcal{G}_p$, the parameter space of the information sharing game $\mathcal{G}_{pq}$ has another dimension, the social fidelity q. We search for the Nash equilibrium strategy profiles and fitness maximizers of $\mathcal{G}_{pq}$ over the parameter space $(p, q, \kappa) \in [1/2, 1]^2 \times (0, \infty)$. First, we observe that unlike $\mathcal{G}_p$, the coordination structure in the new set of strategies $\mathcal{S}^{16}$ is not preserved in $\mathcal{G}_{pq}$. That is, not all Nash equilibria are necessarily symmetric strategy profiles for all $(p, q) \in [1/2, 1]^2$. This means we cannot restrict our search for Nash equilibria and fitness maximizers to symmetric strategy profiles. Indeed, we find that the asymmetric strategy profile (s_1, s_2) with $s_1 = s_{\mathrm{FC}} = [A, A, B, B]$ and $s_2 = [A, B, A, B]$ is a Nash equilibrium in a region of low sensing fidelity p and high sharing q. We refer the reader to Appendix 2.C where we derive its equilibrium region.

There are 136 unique strategy profiles (by symmetry). In Figure 2.5, we display the multiplicity of Nash equilibria across the range of fidelity parameters p and q. Here, we do not count OA and OB because they are always Nash equilibria. We find there is an abundance of Nash equilibria in the upper left region of the parameter space $(p, q) \in [1/2, 1]^2$, where private information is unreliable but information sharing has high fidelity. This suggests reliable social cues are utilized as a means to coordinate (see Section 2.3 for further discussion).

The strategy profiles that maximize fitness are shown in Figure 2.6 in the parameter space $(p, q) \in [1/2, 1]^2$ and for cross-sections of κ. We note that the set of fitness maximizers consist only of symmetric strategy profiles. We find that OA, OB, and FC are fitness maximizers, along with a new type of strategy we term "Majority Logic" (ML). When environment A is favored ($\kappa > 1$), ML_A appears as a fitness maximizer,

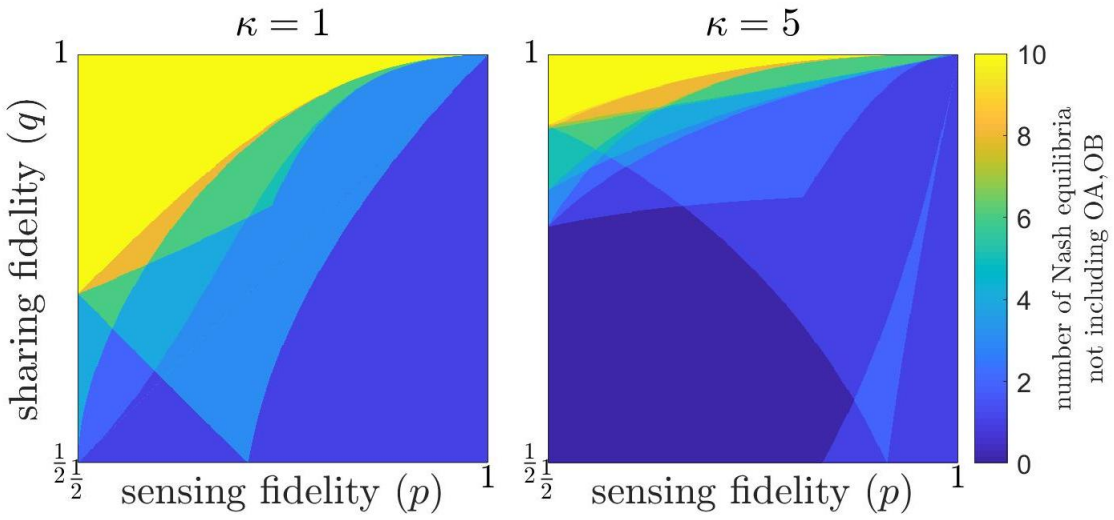


Figure 2.5: The number of Nash equilibria not including OA or OB, that exist across fidelity values $(p, q) \in [1/2, 1]^2$, for $\kappa = 1$ (Left) and $\kappa = 10$ (Right) in the game $\mathcal{G}_{pq}$. The data is numerically calculated by sweeping through $(p, q) \in [1/2, 1]^2$ in a uniform grid of spacing .001, and exhaustively verifying whether each 136 strategy profile satisfies its Nash condition. We verify whether a strategy profile is a Nash equilibrium or not by using a simplified expression of the Nash condition that is amenable to numerical evaluation, given in Appendix 2.C.

and similarly ML_B when $\kappa < 1$. These strategies are written

$$s_{\text{ML}_A} \equiv [A, A, A, B]$$

$$s_{\text{ML}_B} \equiv [A, B, B, B].$$

Recall that in our strategy vector notation, $[A, A, A, B]$ is the strategy where the player chooses action A when $y_i = (e_A, e_A)$, action A when $y_i = (e_A, e_B)$, action A when $y_i = (e_B, e_A)$, and action B when $y_i = (e_B, e_B)$. The majority logic strategies ML_A and ML_B were not previously available in $\mathcal{G}_p$. Their fitnesses increase quadratically in q, and their functional forms given in Appendix 2.A.

The notation ML_A,ML_B differentiates which action is assigned, A or B, to the cues $y_i = (e_A, e_B)$ and (e_B, e_A), i.e. the middle two entries in the strategy vector. We will write just "ML" when generally speaking of the majority logic strategies (e.g. when $\kappa = 1$). Figure 2.9 plots the fitness surfaces of ML and FC.

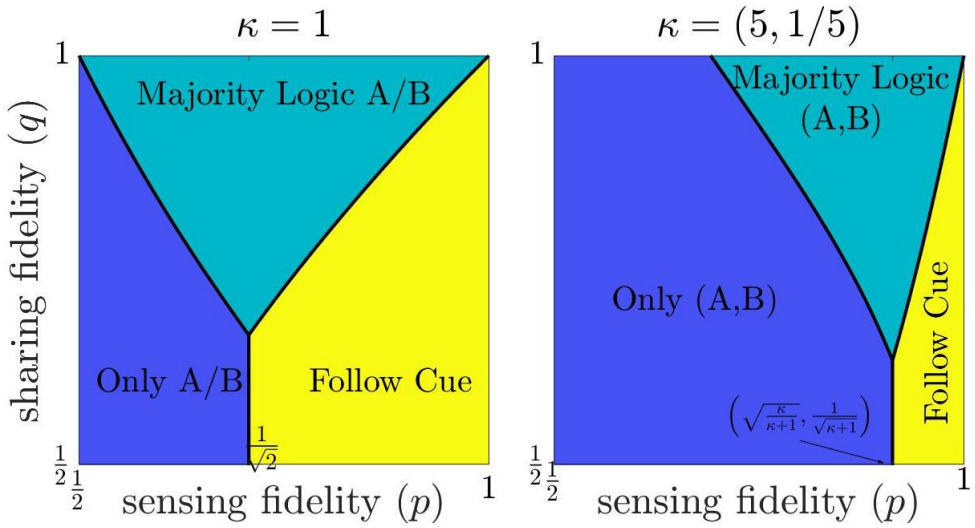


Figure 2.6: Three classes of strategies emerge as the fitness maximizers in $\mathcal{G}_{pq}$: 1) the pure strategies OA and OB, 2) follow cue (FC), and 3) majority logic (ML_A and ML_B). The regions are drawn by first sweeping $(p, q) \in [1/2, 1]^2$ in a uniform grid with spacing .001, and exhaustively searching all 136 strategy profiles for the fitness maximizer. Each grid cell is then filled with the unique color corresponding to the fitness-maximizing strategy profile. Upon observing the emergence of the three strategy classes, the boundaries are analytically solved by equating their fitnesses. Hence, results are accurate within a spacing tolerance of .001. (Left) When $\kappa = 1$, players do not prefer any environment over the other. Hence, OA and OB give the same fitness value and we indicate this by OA/OB (similarly for ML_A/ML_B). (Right) When $\kappa \neq 1$, the region boundaries are the same for both 5 and $\frac{1}{5}$ but the fitness maximizer is either OA or OB, depending on whether κ is greater or less than 1, respectively (similarly, either ML_A or ML_B).

The region where the ML strategies maximize fitness is characterized by intermediate sensing fidelity p and high sharing fidelity q. The shared cues β_i are highly accurate so each player will have reliable knowledge of the other's private cue. Our interpretation of why ML thrives in this regime is as follows. With reliable information sharing, players can detect when their private cues agree and when they disagree. Consider $s_{\mathrm{ML}_A} = [A, A, A, B]$. When they believe both of their private cues match the correct environment (e.g. $y_1 = y_2 = (e_A, e_A)$ w.p. p^2q^2), they choose the correct action. When they believe their private cues disagree with each other (e.g. $y_1 = (e_A, e_B)$ and $y_2 = (e_B, e_A)$ w.p. $p(1-p)q^2$), they decide to coordinate on the predetermined action A, whether or not it is the correct action. These situations will occur far more frequently than players obtaining polarized beliefs (e.g. $y_1 = (e_A, e_A)$ and $y_2 = (e_B, e_B)$ w.p. $p(1-p)(1-q)^2$), where they miscoordinate actions.

The s_{ML_A} strategy vector $[A, A, A, B]$ differs from $s_{\mathrm{FC}} = [A, A, B, B]$ and $s_{\mathrm{OA}} = [A, A, A, A]$ by only one bit, whereas s_{FC} differs from s_{OA} by two bits (similarly for s_{ML_B} and s_{OB}). Thus, we interpret s_{ML} to be a hybrid of s_{FC} and s_{OA}. It acts as an estimator of the environment when the posterior belief on the environmental state is very high, i.e. when $y_i = (e_A, e_A)$ or (e_B, e_B). This is the reasoning for the term "Majority Logic". Players act upon their inference of the environmental state only when their private information is validated by social information. Otherwise, when $y_i = (e_A, e_B)$ or (e_B, e_A), the posterior is not as high and the player disregards its information altogether, blindly playing action A.

2.2.3 The fitness value of information sharing

Majority logic stands alone among all possible strategies in $\mathcal{G}_{pq}$ that can outperform the fitness maximizers of $\mathcal{G}_p$ (OA, OB, and FC). However, recall from Figure 2.6 (Left, for example) that it requires the sharing fidelity q to be sufficiently high for a fixed sensing fidelity p. This critical threshold value, which we call $q_c(p)$, is the value of q

above which the majority logic strategy is the fitness maximizer. Hence, the values of $q_c(p)$ are parameterized by the boundary lines that separate the ML strategies from OA/OB and FC in Figure 2.6. Full parameterizations of $q_c(p)$ are given in Appendix 2.D.

The critical thresholds $q_c(p)$ when the ML strategies outperform the optimal strategies of $\mathcal{G}_p$ suggests there is no value for players to share signals unless sharing fidelity is sufficiently high, $q > q_c(p)$. Higher sharing fidelity is needed for extreme values of p : $q_c(p)$ increases up to 1 as p decreases towards 1/2, as well as when p increases towards 1. When p is near 1, players prefer FC because they are able to independently detect the correct environment with very high probability and act accordingly. When p is near 1/2, private cues are effectively random because they contain no information about E. Consequently, the shared signals are also effectively random. Players then receive any of the four signals y_i with equal probability. By employing ML in this regime, players will miscoordinate more often than they would if they committed to OA or OB.

We are also interested in quantifying the fitness benefit of sharing signals over no sharing. Let

$$F_p(\kappa) \equiv \max_{\mathrm{X=OA,OB,FC}} f_p(\mathrm{X})$$

be the maximum fitness attainable in $\mathcal{G}_p$ at (p, κ), and let

$$F_{pq}(\kappa) \equiv \max_{\mathrm{X=OA,OB,FC,ML}_A,\mathrm{ML}_B} f_{pq}(\mathrm{X})$$

be the maximum fitness attainable in $\mathcal{G}_{pq}$ at (p, q, κ). Then we define the *fitness value of information sharing* as

$$V(p, q, \kappa) \equiv \frac{F_{pq}(\kappa)}{F_p(\kappa)},$$

which gives the ratio of the maximum fitness in $\mathcal{G}_{pq}$ to the maximum fitness in $\mathcal{G}_p$ with the parameters (p, q, κ). By definition, $V(p, q, \kappa) = 1$ for $q \leq q_c(p)$. In other words, there is no fitness benefit to sharing signals when q does not exceed the threshold

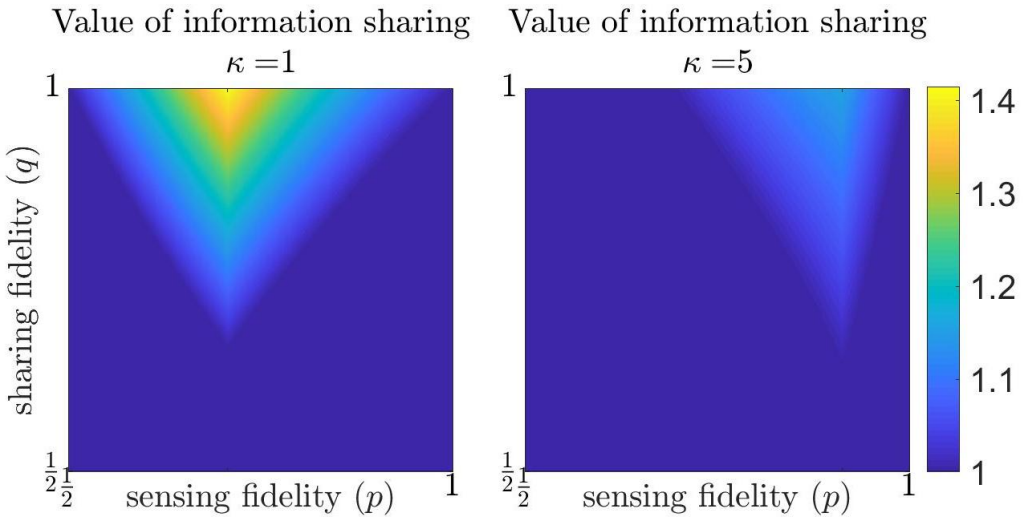


Figure 2.7: The fitness value of information sharing $V(p, q, \kappa)$. (Left) When $\kappa = 1$, the maximum value of V is attained at $(p = 1/\sqrt{2}, q = 1, \kappa = 1)$, where information sharing leads to approximately a 41% increase in fitness over no sharing. (Right) When $\kappa = 5$, information sharing leads to approximately 15% increase in fitness at best $(p = \sqrt{\frac{\kappa}{1+\kappa}}, q = 1)$.

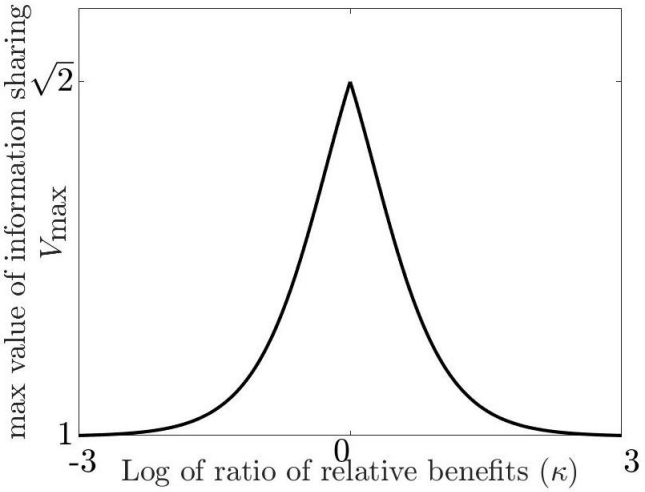


Figure 2.8: The maximum value of information sharing $V_{\max}(\kappa)$, which is the value of $V(p, q, \kappa)$ attained at $q = 1$ and $p = 1/\sqrt{1+\kappa}$ when $\kappa < 1$ and $p = \sqrt{\kappa/(1+\kappa)}$ when $\kappa \geq 1$. The peak of fitness improvement occurs when no environment is favored over the other $(\kappa = 1)$, where there is an improvement ratio of $\sqrt{2}$. The value for $V_{\max}$ is $f_{pq}(\mathrm{ML}_B)/c_B$ when $\kappa < 1$, and $f_{pq}(\mathrm{ML}_A)/c_A$ when $\kappa > 1$ (they coincide when $\kappa = 1$). The fitness improvement ratio degrades as $\kappa \to 0$ or ∞, i.e. when one environment is favored over the other.

$q_c(p)$. A contour map of V is shown in Figure 2.7. When $\kappa = 1$ (Figure 2.7, Left), V is maximized at $p = 1/\sqrt{2}$ and $q = 1$, giving $V_{\max} = \sqrt{2}$ (use Eq. (2.9) and divide by c_B). Thus, information sharing improves fitness by approximately 41% when the environment fluctuates frequently. For $\kappa > 1$, it is maximized at $q = 1$, $p = \sqrt{\kappa/(1+\kappa)}$ (Figure 2.7, Right). The general form of $V_{\max}(\kappa)$ is the following piecewise continuous function,

$$V_{\max}(\kappa) = \begin{cases} \sqrt{\frac{\kappa}{\kappa+1}}\left(2 - \sqrt{\frac{\kappa}{\kappa+1}}(1-\kappa^{-1})\right) & \text{if } \kappa > 1 \\ \sqrt{2} & \text{if } \kappa = 1 \\ \frac{1}{\sqrt{\kappa+1}}\left(2 - \frac{1}{\sqrt{\kappa+1}}(1-\kappa)\right) & \text{if } 0 < \kappa < 1 \end{cases}$$

and is plotted in Figure 2.8. As $\kappa \to 0$ or ∞, the improvement ratio $V_{\max}$ degrades as one environment becomes favored over the other. In these extreme scenarios, either OA or OB become optimal for increasingly larger regions.

We also note that $V(p, q, \kappa)$ increases quadratically in q for $q > q_c(p)$. Hence, the rate at which V increases with respect to q, $\frac{dV}{dq}$, increases linearly in q. This is due to a majority logic strategy being the fitness maximizer in this region. We also note that $V(p, q, \kappa)$ decreases as p deviates away from the intermediate value $\sqrt{\kappa/(1+\kappa)}$. At the extremes $p = 1/2$ and $p = 1$, the strategies OA and FC begin to dominate, respectively, as the line $q_c(p)$ tapers to 1.

2.3 Discussion

In this chapter, we have presented a two-player, two action coordination game in a fluctuating environment where players independently sense private cues from the environment and share their cues with each other. In our analysis, we found the optimal strategies that promote coordination across all levels of sensing and sharing fidelities (Figures 2.4 and 2.6). When individual sensing is very reliable, there is no need to share signals because players can accurately infer the environmental state

independently. When sensing is unreliable, players prefer to ignore their information altogether and always commit to a single action. The "majority logic" strategy is optimal when private information has intermediate fidelity and social information has high fidelity. This strategy highlights the importance of information sharing because they are the only optimal strategies that utilize the shared social cues. Because these optimal strategies are derived from a first-principles approach, their appearance in our model offers insight into the mechanisms that maintain group behaviors.

The Majority Logic strategies strike a balance between when to use information as a predictor of the environment and when to use information as a means to coordinate. Essentially, it allows players to coordinate when their individual inferences (private information) about the environment conflict, and to choose the correct action when they agree. The interplay of private and social information in our model draws similarities to Condorcet's jury theorem, where in [52], states that good reliability of private information is a requirement to effectively make group decisions based on social information in the large population limit. In their model, individuals disregard social information when private information is poor. When it is more accurate, individuals rely more on social information to make a decision. In our model, these two situations are akin to the fitness maximizing regions of OA and OB for poor sensing fidelity, and the region where ML thrives, respectively. These results also corroborate with controlled experimental lab work conducted on nine-spined sticklebacks [108]. When private experience about foraging sites was 100% reliable, sticklebacks based foraging decisions only on private information. When it was less reliable, they followed social cues instead. Without perfect sensing capabilities, organisms need to rely on social information to survive [108, 52, 83, 5, 64].

In the region of unreliable private information and reliable social information (Fig 2.5, right), there is an abundance of Nash equilibria (ten) that are suboptimal to the OA and OB strategies. This suggests that in the context of coordination games,

social cues serve as a coordination device rather than as an additional source of information about the environment. However, the strategies included in this region, which includes majority logic and its variants, will often coordinate on the wrong action. This is because the social cues carry no information about the environment, as the private cues are themselves uninformative. The best an individual can do to infer the environment is simply to guess. However, if both guess independently, they will only coordinate 1/4 of the time. If they play OA, they are guaranteed to coordinate for the fraction of time the environment spends in state e_A. Hence, committing to a single action corresponding to the most frequent environment (OA or OB) is the best the group can do when information about the environment is poor. Therefore in principle, information sharing is useful only when the information that is being shared is itself reliable.

Our game-theoretic model portrays situations where communicating individuals must coordinate behaviors in uncertain fluctuating environments. These situations pervade collective behaviors in groups of organisms across the animal kingdom from schools of fish to bacterial colonies [**?**]. The information flow in our model is particularly inspired by quorum sensing in bacterial populations. Such a communication system enables bacteria to display complex social behaviors [120]. Therefore, game theory is a natural framework in which to study microbial decision-making to consolidate experimental understanding of the phenomena.

Several questions remain in our study. Our work has not yet considered players with memory of past signals and actions. Repeated games that endow players with these reasoning capabilities, for example the iterated prisoner's dilemma (IPD), gives rise to new classes of strategies that offer insights into the evolution of cooperation [88, 107]. Such an extension of our current formulation would give players, through an expanded space of memory-dependent strategies, the ability to anticipate environmental switches and the intentions of others. These considerations warrant

future research. Furthermore, better quality signalling entails increasing fitness costs [15, 16, 104, 61, 47]. Our work has not yet investigated such evolutionary tradeoffs. Instead, we have presented a systematic, centralized analysis of the optimal strategies given a fixed, costless communication system. Consequently, we have yet to address whether the optimal strategies identified are stable in an evolutionary sense, with or without costly signalling.

Moving forward, evolutionary dynamics can be applied to our formulation in the context of population games [99]. Population games are a framework to describe the interactions between a well-mixed, continuous mass of agents that select from the same set of strategies. In the population, interactions between agents are probabilistic and pair-wise, which allows two-player normal form games to be represented as population games. Such a framework works well with our aim to describe a population of organisms that can adopt a variety of communication-based strategies, *e.g.*, quorum sensing bacteria. Furthermore, due to the potential structure of our games $\mathcal{G}_p$ and $\mathcal{G}_{pq}$ (see Remark 2.3), they admit evolutionary dynamics that have certain stability guarantees [98, 99]. Embedding our model into population games will be necessary to identify which local maximizers of average fitness are likely to be reached. In doing so, we hope that our model encourages the integration of social interactions and communication into efforts to understand coordination, cooperation, and conflict in complex environments.

Appendix

For notational convenience in the following Appendices, we abuse notation of strategy vectors s_i by assigning the numeric value $s_i(y_i) = 0$ if action A was assigned, and $s_i(y_i) = 1$ if B was assigned (y_i is either α_i or (α_i, β_i)). For instance, the strategy vector $s_{\text{FC}}^\top = [A, B]$ corresponds to the vector $s_{\text{FC}}^\top = [0, 1]$ whenever computations are needed. We will adopt the $0, 1$ notation for strategy vectors throughout this supplementary document.

2.A Calculation of fitness functions f_p and f_{pq}

Here we calculate the fitnesses $f_p(s_1, s_2)$ and $f_{pq}(s_1, s_2)$ explicitly in terms of p, q, and the strategy vectors s_1 and s_2.

The fitness f_p

In the game $\mathcal{G}_p$, recall the average long-term fitness is (eq. (2.3) in the main text)

$$f_p(s_1, s_2) = \sum_{\alpha_1, \alpha_2, E} \pi_p(\alpha_1, \alpha_2, E) U(s_1(\alpha_1), s_2(\alpha_2), E). \tag{2.7}$$

It is a sum of eight terms since $(\alpha_1, \alpha_2, E) \in \{E_A, E_B\}^3$. We define $[Q_A^p]_{\alpha_1, \alpha_2} = P(\alpha_1, \alpha_2 | E_A)$ as the matrix of conditional probabilities on $E = E_A$, and similarly for Q_B for $E = E_B$,

$$Q_A^p \equiv \begin{bmatrix} p^2 & p\bar{p} \\ p\bar{p} & \bar{p}^2 \end{bmatrix}, \quad Q_B^p \equiv \begin{bmatrix} \bar{p}^2 & p\bar{p} \\ p\bar{p} & p^2 \end{bmatrix}$$

where $\bar{p} = 1 - p$.

The only nonzero terms in (2.7) occur when $U(s_1(\alpha_i), s_2(\alpha_j), E_A) = b_A$, which happens if and only if $s_1(\alpha_i) = s_2(\alpha_j) = 0$ (similarly, b_B for $E = E_B$). Therefore, the fitness can be written

$$f_p(s_1, s_2) = c_A(\mathbf{1} - s_1)^\top Q_A^p(\mathbf{1} - s_2) + c_B s_1^\top Q_B^p s_2$$

where recall that $c_A = b_A \frac{v_{BA}}{v_{BA}+v_{AB}}$, $c_B = b_B \frac{v_{AB}}{v_{BA}+v_{AB}}$, and we denote $\mathbf{1}$ as the vector $[1,1]^\top$. The expression can alternatively be written in a bilinear form,

$$f_p(s_1, s_2) = s_1^\top Q^p s_2 - c_A L_p^\top (s_1 + s_2) + c_A. \tag{2.8}$$

where

$$Q^p \equiv c_A Q_A^p + c_B Q_B^p$$

$$L_p^\top \equiv [p, \bar{p}].$$

The fitness f_{pq}

In the game $\mathcal{G}_{pq}$, recall the average long-term fitness is (eq. (2.5) in the main text)

$$f_{pq}(s_1, s_2) \equiv \sum_{y_1, y_2, E} \pi_{pq}(y_1, y_2, E) U(s_1(y_1), s_2(y_2), E).$$

The expressions for f_{pq} are derived in the same manner as f_p (eq. (2.8)). However, we need to define the new matrices Q_A and Q_B whose entries are the conditional probabilities $P(y_1, y_2|E_k)$, $k = A, B$, and where $y_i = (\alpha_i, \beta_i) \in \{E_A, E_B\}^2$.

$$Q_A^{pq} = \left[\begin{array}{cc|cc} p^2q^2 & p^2q\bar{q} & p\bar{p}q\bar{q} & p\bar{p}\bar{q}^2 \\ p^2q\bar{q} & p^2\bar{q}^2 & p\bar{p}q^2 & p\bar{p}q\bar{q} \\ \hline p\bar{p}q\bar{q} & p\bar{p}q^2 & \bar{p}^2\bar{q}^2 & \bar{p}^2q\bar{q} \\ p\bar{p}\bar{q}^2 & p\bar{p}q\bar{q} & \bar{p}^2q\bar{q} & \bar{p}^2q^2 \end{array}\right], \quad Q_B^{pq} = \left[\begin{array}{cc|cc} \bar{p}^2q^2 & \bar{p}^2q\bar{q} & p\bar{p}q\bar{q} & p\bar{p}\bar{q}^2 \\ \bar{p}^2q\bar{q} & \bar{p}^2\bar{q}^2 & p\bar{p}q^2 & p\bar{p}q\bar{q} \\ \hline p\bar{p}q\bar{q} & p\bar{p}q^2 & p^2\bar{q}^2 & p^2q\bar{q} \\ p\bar{p}\bar{q}^2 & p\bar{p}q\bar{q} & p^2q\bar{q} & p^2q^2 \end{array}\right]$$

Then $f_{pq}(s_1, s_2)$, where s_1, s_2 are four-vectors of ones and zeros, can be written

$$f_{pq}(s_1, s_2) = c_A(\mathbf{1} - s_1)^\top Q_A^{pq}(\mathbf{1} - s_2) + c_B s_1^\top Q_B^{pq} s_2$$

where $\mathbf{1} = [1,1,1,1]^\top$. An alternative bilinear form is

$$f_{pq}(s_1, s_2) = s_1^\top Q^{pq} s_2 - c_A L_{pq}^\top (s_1 + s_2) + c_A.$$

where

$$Q^{pq} \equiv c_A Q_A^{pq} + c_B Q_B^{pq}$$

$$L_{pq}^\top \equiv [p(pq + \bar{p}\bar{q}), p(p\bar{q} + \bar{p}q), \bar{p}(pq + \bar{p}\bar{q}), \bar{p}(p\bar{q} + \bar{p}q)]$$

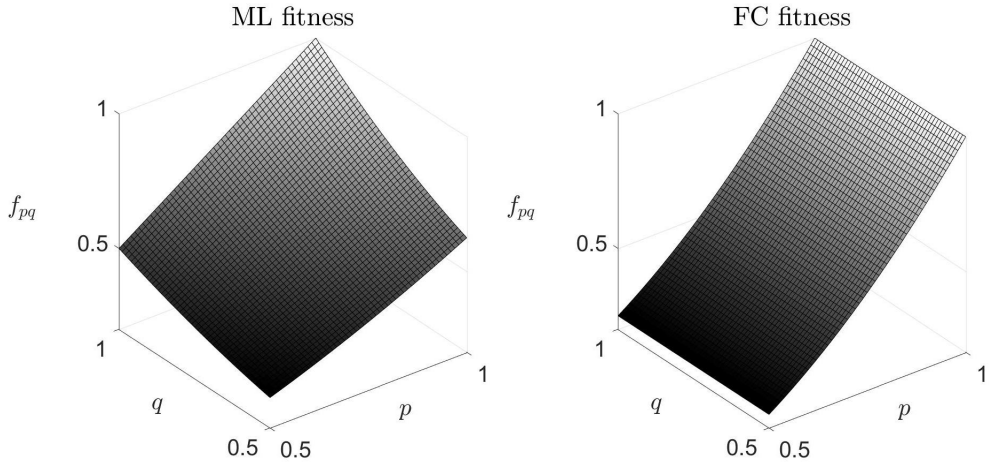


Figure 2.9: Fitnesses of the ML (Left) and FC (Right) strategy profiles for $\kappa = 1$ ($f_{pq}(\mathrm{ML}_A) = f_{pq}(\mathrm{ML}_B)$). In these contours, $b_A = b_B = 1$ and $v_{AB} = v_{BA}$, so $c_A = c_B = 0.5$ and the environments spend equal amounts of time in the long run. Hence, the fitness values on the z axis portray the fraction of time the players coordinate on the correct action. The fitnesses are $f_{pq}(\mathrm{ML}) = (1 - 2p\bar{p})(q^2 + 1) + 2\bar{p}(2p - 1)q$ and $f_{pq}(\mathrm{FC}) = p^2$.

The fitness of FC is

$$f_{pq}(\mathrm{FC}) = (c_A + c_B)p^2$$

The fitnesses of ML_A and ML_B are

$$\begin{aligned} f_{pq}(\mathrm{ML}_A) &= f_{pq}(s_{\mathrm{ML}_A}, s_{\mathrm{ML}_A}) = (c_A\bar{p}^2 + c_Bp^2)q^2 + 2c_A\bar{p}(2p - 1)q + c_A(1 - 2p\bar{p}) \\ f_{pq}(\mathrm{ML}_B) &= f_{pq}(s_{\mathrm{ML}_B}, s_{\mathrm{ML}_B}) = (c_Ap^2 + c_B\bar{p}^2)q^2 + 2c_B\bar{p}(2p - 1)q + c_B(1 - 2p\bar{p}). \end{aligned} \tag{2.9}$$

They increase quadratically in q. Figure 2.9 plots the fitness surfaces of FC and ML.

2.B Proof of Proposition 2.1

We reproduce the statement from Section 2.2.1.

For $p \in [1/2, 1]$, all Nash equilibria and fitness maximizers of $\mathcal{G}_p$ are necessarily symmetric strategy profiles, i.e. $s_1^* = s_2^*$.

Proof. This fact is proven by exhaustion. For each of the six asymmetric strategy

profiles, we show there is at least one player that can imitate the other's strategy to gain fitness. This implies Nash equilibria and fitness maximizers are necessarily symmetric strategy profiles. The fitnesses of these six profiles are given by the entries below (or above) the diagonal of the payoff matrix (2.4), reproduced here.

$$\begin{array}{c} \\ s_{\mathrm{OA}} \\ s_{\mathrm{FC}} \\ s_{\overline{\mathrm{FC}}} \\ s_{\mathrm{OB}} \end{array} \begin{array}{c} \begin{array}{cccc} s_{\mathrm{OA}} & s_{\mathrm{FC}} & s_{\overline{\mathrm{FC}}} & s_{\mathrm{OB}} \end{array} \\ \begin{bmatrix} c_A & c_A p & c_A \bar{p} & 0 \\ c_A p & (c_A + c_B)p^2 & (c_A + c_B)p\bar{p} & c_B p \\ c_A \bar{p} & (c_A + c_B)p\bar{p} & (c_A + c_B)\bar{p}^2 & c_B \bar{p} \\ 0 & c_B p & c_B \bar{p} & c_B \end{bmatrix} \end{array}$$

1. $f_p(s_{\mathrm{OA}}, s_{\mathrm{FC}}) = c_A p$. If the s_{FC} player switches to s_{OA}, the fitness becomes $f_p(s_{\mathrm{OA}}, s_{\mathrm{OA}}) = c_A > c_A p$ iff $p \neq 1$. If the s_{OA} player switches to s_{FC}, the fitness becomes $f_p(s_{\mathrm{FC}}, s_{\mathrm{FC}}) = (c_A + c_B)p^2 > c_A p$ iff $p > c_A/(c_A + c_B)$, which is satisfied at $p = 1$.

2. $f_p(s_{\mathrm{OA}}, s_{\overline{\mathrm{FC}}}) = c_A \bar{p}$. If the $s_{\overline{\mathrm{FC}}}$ player switches to s_{OA}, the fitness becomes $f_p(s_{\mathrm{OA}}, s_{\mathrm{OA}}) = c_A > c_B \bar{p}$.

3. $f_p(s_{\mathrm{OA}}, s_{\mathrm{OB}}) = 0$. If the s_{OB} player switches to s_{OA}, the fitness becomes $f_p(s_{\mathrm{OA}}, s_{\mathrm{OA}}) = c_A > 0$.

4. $f_p(s_{\mathrm{FC}}, s_{\overline{\mathrm{FC}}}) = (c_A + c_B)p\bar{p}$. If the $s_{\overline{\mathrm{FC}}}$ player switches to s_{FC}, the fitness becomes $f_p(s_{\mathrm{FC}}, s_{\mathrm{FC}}) = (c_A + c_B)p^2 > (c_A + c_B)p\bar{p}$.

5. $f_p(s_{\mathrm{FC}}, s_{\mathrm{OB}}) = c_B p$. If the s_{FC} player switches to s_{OB}, the fitness becomes $f_p(s_{\mathrm{OB}}, s_{\mathrm{OB}}) = c_B > c_B p$ if and only if $p \neq 1$. If the s_{OB} player switches to s_{FC}, the fitness becomes $f_p(s_{\mathrm{FC}}, s_{\mathrm{FC}}) = (c_A + c_B)p^2 > c_B p$ iff $p > c_B/(c_A + c_B)$, which is satisfied at $p = 1$.

6. $f_p(s_{\overline{\mathrm{FC}}}, s_{\mathrm{OB}}) = c_B \bar{p}$. If the $s_{\overline{\mathrm{FC}}}$ player switches to s_{OB}, the fitness becomes $f_p(s_{\mathrm{OB}}, s_{\mathrm{OB}}) = c_B > c_B \bar{p}$.

■

2.C Nash equilibrium conditions

The game $\mathcal{G}_p$ Due to the statement in Appendix 2.B, we only need to consider the four symmetric strategy profiles to find the Nash equilibria of the game $\mathcal{G}_p$. We give an expression to check whether a symmetric strategy profile is a Nash equilibrium.

Theorem 2.1. *The strategy profile (s^*, s^*) is a (strict) Nash equilibrium in $\mathcal{G}_p$ if*

$$diag(2s^* - \mathbf{1})(Q^p s^* - c_A L_p) \succ 0$$

where $\mathbf{1} = [1, 1]^\top$ and $\succ$ denotes element-wise strict inequality.

Proof. We define player i's best response s_i^* to player j's strategy s_j as

$$s_i^* = \mathrm{BR}_i(s_j) \equiv \arg\max_{s_i} f_p(s_i, s_j)$$

Since $f_p(s_i, s_j)$ is bilinear in s_i and s_j, the best-response can be written

$$s_i^*(E_k) = \begin{cases} 0 \text{ if } [Q^p s_j - c_A L_p]_{E_k} < 0 \\ 1 \text{ if } [Q^p s_j - c_A L_p]_{E_k} \geq 0 \end{cases}, \ k = A, B$$

When the condition above is met with equality, player i is indifferent to choosing 0 or 1 at $s_i^*(E_k)$. For convention, we define the best-response to choose 1 in this case. The definition of a Nash equilibrium (s_1, s_2) states

$$s_1 \in \mathrm{BR}_1(s_2) \text{ and } s_2 \in \mathrm{BR}_2(s_1)$$

Since we can consider only symmetric strategy profiles, the condition simplifies to $s \in \mathrm{BR}(s)$ ■

The Nash equilibrium regions for each of the four strategy profiles are derived by applying Theorem 2.1, which gives two inequality conditions in terms of p and κ (after normalizing by dividing c_B). The fitness maximizer regions are determined by

comparing fitnesses for each strategy. For instance, the region where OA maximizes fitness is derived by setting $f_p(\text{OA}) > f_p(\text{X})$ for each X = FC, $\overline{\text{FC}}$, and OB. Table 3 lists the properties of all symmetric strategy profiles in $\mathcal{G}_p$.

Table 3: Properties for the symmetric strategy profiles of $\mathcal{G}_p$.

Profile (s, s)	vector form	$f_p(s, s)$	NE region (p, κ)	fitness max region
OA	$[0, 0]^\top$	c_A	everywhere	$\kappa > \{p^2/(1-p^2), 1\}$
FC	$[0, 1]^\top$	$(c_A + c_B)p^2$	$\kappa > \bar{p}/p,\ \kappa < p/\bar{p}$	$\kappa < p^2/(1-p^2), \kappa > 1/p^2 - 1$
$\overline{\text{FC}}$	$[1, 0]^\top$	$(c_A + c_B)\bar{p}^2$	nowhere	nowhere
OB	$[1, 1]^\top$	c_B	everywhere	$\kappa < \{1/p^2 - 1, 1\}$

The game with information sharing $\mathcal{G}_{pq}$

The Nash equilibrium condition for the game $\mathcal{G}_{pq}$ is stated here.

Theorem 2.2. *The strategy profile (s_1^*, s_2^*) is a (strict) Nash equilibrium in $\mathcal{G}_{pq}$ if*

$$diag(2s_1^* - \mathbf{1})(Q^{pq}s_2^* - c_A L_{pq}) \succ 0$$

$$diag(2s_2^* - \mathbf{1})(Q^{pq}s_1^* - c_A L_{pq}) \succ 0$$

Proof. Similar arguments as in Theorem 2.1. ■

Corollary 2.1. *In $\mathcal{G}_{pq}$, FC is a Nash equilibrium for*

$$q < \frac{\kappa p}{\bar{p} + \kappa p}, \qquad \text{if } \kappa < 1$$

$$q < \frac{p}{p + \kappa \bar{p}}, \qquad \text{if } \kappa > 1$$

$$q < p, \qquad \text{if } \kappa = 1$$

The region is shown in Figure 2.10.

Proof. Recall $s_{\text{FC}} = [0, 0, 1, 1]^\top$. From Theorem 2.2, the NE condition for $(s_{\text{FC}}, s_{\text{FC}})$ is

$$\text{diag}(2s_{\text{FC}} - \mathbf{1})(Q^{pq}s_{\text{FC}} - c_A L_{pq}) \succ 0$$

After dividing by c_B and p, we obtain four inequality conditions

$$\begin{aligned} -\kappa pq + \bar{p}\bar{q} &< 0 \\ -\kappa p\bar{q} + \bar{p}q &< 0 \\ -\kappa \bar{p}q + p\bar{q} &> 0 \\ -\kappa \bar{p}\bar{q} + pq &> 0 \end{aligned} \tag{2.10}$$

If $\kappa = 1$, the first and fourth inequality conditions above are the same, and the second and third are the same. This reduces (2.10) to two conditions - 1) $pq > \bar{p}\bar{q}$ and 2) $p\bar{q} > \bar{p}q$. Here, 1) is always true, and 2) reduces to $p > q$. For general $\kappa \neq 1$, we obtain

$$\max\left\{\frac{\bar{p}}{\bar{p}+p\kappa}, \frac{\kappa\bar{p}}{p+\bar{p}\kappa}\right\} < q < \min\left\{\frac{\kappa p}{\bar{p}+p\kappa}, \frac{p}{p+\bar{p}\kappa}\right\}$$

If $\kappa < 1$, it is $\frac{\bar{p}}{\bar{p}+p\kappa} < q < \frac{\kappa p}{\bar{p}+p\kappa}$. If $\kappa > 1$, it is $\frac{\kappa\bar{p}}{p+\bar{p}\kappa} < q <, \frac{p}{p+\bar{p}\kappa}$ ■

Corollary 2.2. *In $\mathcal{G}_{pq}$, ML_A is a Nash equilibrium for*

$$\begin{aligned} p((1+\kappa)q\bar{q} - \kappa) + (\bar{q} - \kappa q) &< 0 \\ p(\kappa(2q-1) - (\kappa+1)q\bar{q}) + (1 - (1+\kappa)q) &< 0 \\ \kappa\bar{p}(\bar{p}q\bar{q} - pq - \bar{p}\bar{q}) + p^2 q\bar{q} &< 0 \\ \kappa\bar{p}(\bar{p}q^2 - pq - \bar{p}\bar{q}) + p^2 q^2 &> 0 \end{aligned}$$

The region is shown in Figure 2.11.

Proof. Apply Theorem 2.2 ■

The Nash Equilibrium region can be derived in a similar manner for ML_B, which coincides with ML_A when $\kappa = 1$.

Corollary 2.3. *In $\mathcal{G}_{pq}$, the asymmetric strategy profile (s_1, s_2) with $s_1 = s_{FC} = [0,0,1,1]$ and $s_2 = [0,1,0,1]$ is a Nash equilibrium for*

$$\begin{aligned} q &> \frac{p}{p+\kappa\bar{p}}, && \text{if } \kappa < 1 \\ q &> \frac{\kappa p}{\bar{p}+\kappa p}, && \text{if } \kappa > 1 \\ q &> p, && \text{if } \kappa = 1 \end{aligned}$$

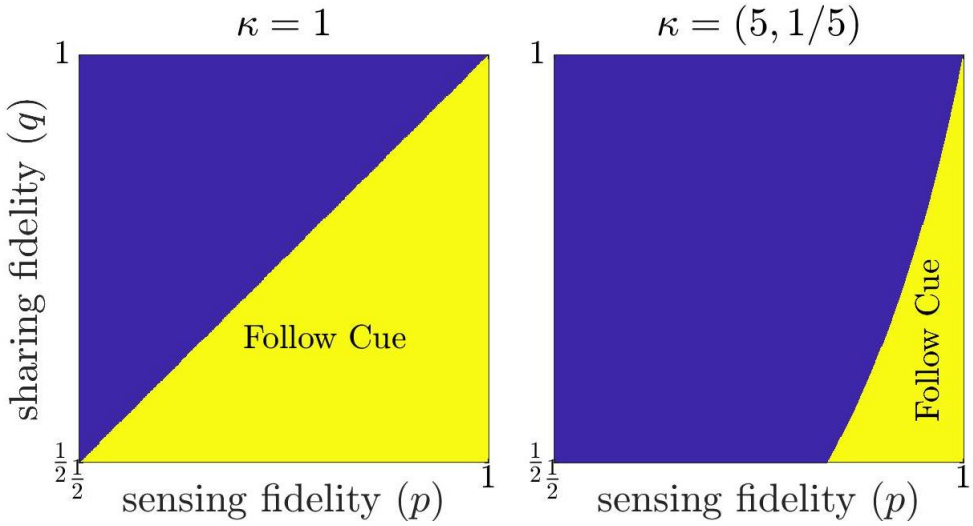


Figure 2.10: Nash equilibrium region for the Follow Cue (FC) strategy profile in the parameter space $(p, q) \in [\frac{1}{2}, 1]^2$, as given by Corollary 2.1.

The s_2 strategy "follows" the social cue sent by player 1. Hence we refer to this strategy profile as "FC-F" (Follow Cue - Follower). The region is shown in Figure 2.12.

Proof. Apply Theorem 2.2 . ■

2.D Parameterizations of critical sharing theshold

We derive the parameterizations for the critical threshold value $q_c(p)$ in the six cases that determine the fitness maximizer in $\mathcal{G}_{pq}$. These curves are the boundaries of the region where the ML strategies are fitness maximizers (see Figure 2.13).

Proposition 2.2. *The critical threshold sensing fidelity q_c is given by the following six parameterizations.*

(a) $q_c = \frac{-2\kappa\bar{p}(2p-1)+\sqrt{(2\kappa\bar{p}(2p-1))^2+8\kappa p\bar{p}(\kappa\bar{p}^2+p^2)}}{2(\kappa\bar{p}^2+p^2)}$ *for* $\frac{\kappa-\sqrt{\kappa}}{\kappa-1} \leq p \leq \sqrt{\frac{\kappa}{\kappa+1}}$, $\kappa > 1$.

(b) $q_c = \frac{-2\kappa\bar{p}(2p-1)+\sqrt{(2\kappa\bar{p}(2p-1))^2-4(\kappa(1-2p\bar{p})-(\kappa+1)p^2)(\kappa\bar{p}^2+p^2)}}{2(\kappa\bar{p}^2+p^2)}$ *for* $\sqrt{\frac{\kappa}{\kappa+1}} \leq p \leq 1$, $\kappa > 1$.

(c) $q_c = \frac{-2\bar{p}(2p-1)+\sqrt{(2\bar{p}(2p-1))^2+8p\bar{p}(\kappa p^2+\bar{p}^2)}}{2(\kappa p^2+\bar{p}^2)}$ *for* $\frac{\sqrt{\kappa^{-1}}-1}{\kappa^{-1}-1} \leq p \leq \sqrt{\frac{1}{\kappa+1}}$, $\kappa < 1$.

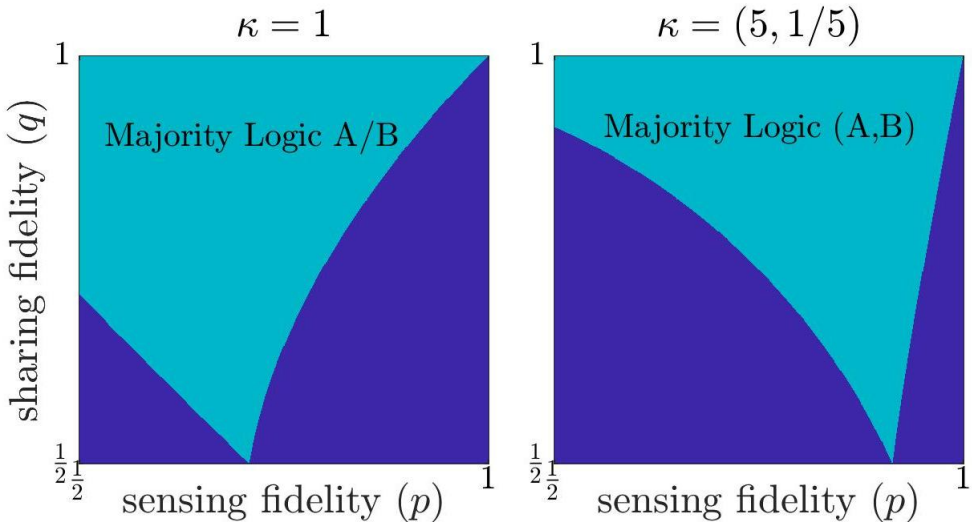


Figure 2.11: Nash equilibrium region for the Majority Logic (ML) strategy profiles in the parameter space $(p, q) \in [\frac{1}{2}, 1]^2$, as given by Corollary 2.2.

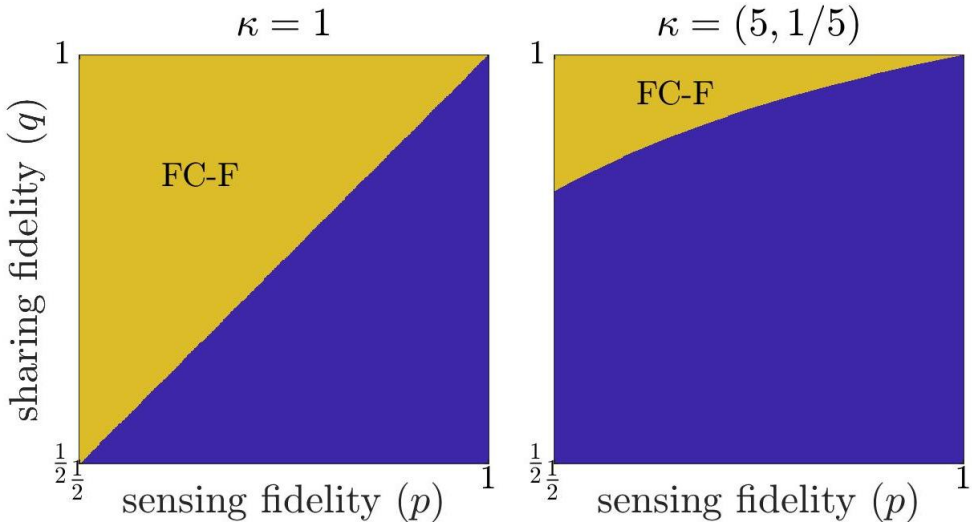


Figure 2.12: Nash equilibrium region for the FC-F strategy profile in the parameter space $(p, q) \in [\frac{1}{2}, 1]^2$, as given by Corollary 2.3.

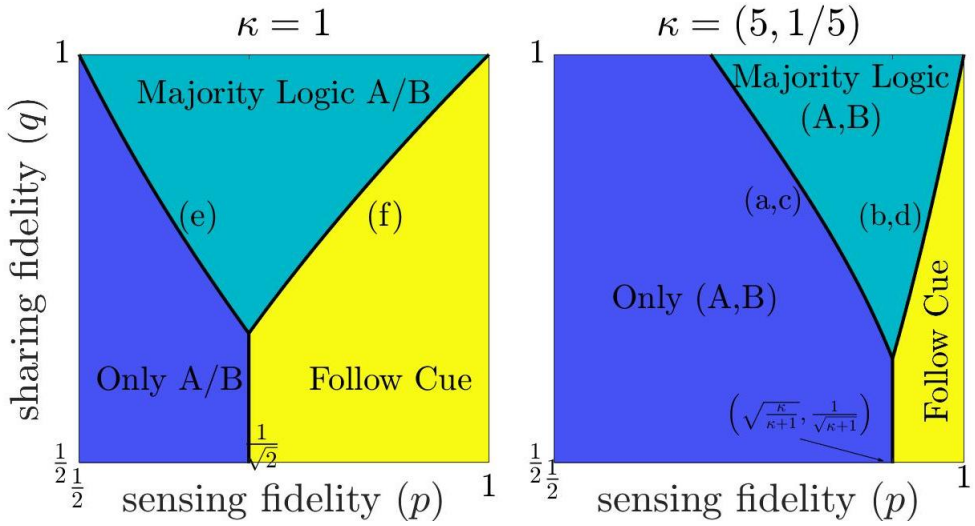


Figure 2.13: Boundary curves labeled (a-f) that define the critical threshold fidelity q_c that separates fitness maximizing regions. The labeled curves are parameterized according to Proposition 2.2, which separate the region where the Majority Logic strategies are fitness maximizers.

(d) $q_c = \frac{-2\bar{p}(2p-1)+\sqrt{(2\bar{p}(2p-1))^2-4((1-2p\bar{p})-(\kappa+1)p^2)(\kappa p^2+\bar{p}^2)}}{2(\kappa p^2+\bar{p}^2)}$ *for* $\frac{\sqrt{\kappa^{-1}}-1}{\kappa^{-1}-1} \leq p \leq 1$, $\kappa < 1$.

(e) $q_c = \frac{-2\bar{p}(2p-1)+\sqrt{(2\bar{p}(2p-1))^2+8p\bar{p}(\bar{p}^2+p^2)}}{2(\bar{p}^2+p^2)}$ *for* $\frac{1}{2} \leq p \leq \sqrt{\frac{1}{2}}$, $\kappa = 1$.

(f) $q_c = \frac{-2\bar{p}(2p-1)+\sqrt{(2\bar{p}(2p-1))^2-4((1-2p\bar{p})-2p^2)(\bar{p}^2+p^2)}}{2(\bar{p}^2+p^2)}$ *for* $\sqrt{\frac{1}{2}} \leq p \leq 1$, $\kappa = 1$.

Proof. Depending on which of the six cases in the main text is given, $f_{pq}(\mathrm{ML}_A)$ or $f_{pq}(\mathrm{ML}_B)$ is equated to the fitness of OA, OB, or FC to derive the expression for q_c. Since q_c must be between 1/2 and 1, there is a lower cutoff for p in cases (a-d) below which the value of q_c is greater than one, and hence not a valid threshold. ■

CHAPTER III

EPIDEMIC AWARENESS

Mathematical models of epidemic spreading over networks characterize how the spatial features of the network structure affects the dynamics of epidemic spread [115, 36, 30, 69, 109, 2, 112]. A simple, yet descriptive formulation for such processes is the susceptible-infected-susceptible (SIS) model, where an individual is either infected or susceptible to infection. A celebrated result in such models identifies a threshold on the infection parameters that determines whether the epidemic eradicates quickly or persists for a long time. Specifically, $\delta/\beta > \lambda_{\max}(A)$ (β is the disease transmission rate, δ the healing rate, and $\lambda_{\max}(A)$ the largest eigenvalue of the network adjacency matrix) is a sufficient condition for the disease to eradicate exponentially fast. The opposite strict inequality is a necessary condition for the disease to persist for a long period of time. The steady state in this regime is often referred to as the endemic or metastable state.

These classical models, however, do not account for human behavioral responses to the epidemic. The nodes in the network interact with their neighbors without restraint, even when the epidemic is highly prevalent. With the widespread availability of social media and news outlets on multiple platforms, individuals may be well-informed about the current state of ongoing epidemics and how to take precautionary measures to avoid getting sick. In the 2009 H1N1 Influenza pandemic, people responded to public service announcements by increasing the frequency of washing hands, staying at home when they or loved ones were sick, or avoiding large public gatherings [105]. In the recent Ebola outbreak in West Africa, a combination of quarantining and sanitary burial methods were shown to significantly reduce the

rate of virus spread [80]. These precautions and social distancing actions effectively limit epidemic spread. Individuals' distancing actions depend on the extent of how informed they are. The dissemination and exchange of information influences the public's behavior, affecting the course of the epidemic itself, in turn affecting the public's behavior again [14, 31, 45]. This feedback loop allows epidemic spreading to coevolve with human social behavior, inducing complex dynamics [116].

Recent research effort has focused on modeling human behavioral elements into existing models of epidemic spreading. A review of the recent literature can be found in [116]. Such models present general challenges for characterizing decentralized and dynamic protection measures and also capture a realistic aspect of disease spread in society. When individuals take social distancing actions based on the level of information they have, they reduce contact with others and the epidemic prevalence reduces significantly [34, 92, 85]. They can become aware of the epidemic by receiving information from their social contacts or from a global broadcast [37, 38, 103]. In certain formulations, awareness raises the epidemic threshold below which the disease dies out fast [102, 101, 97]. Other actions include switching one's contact links, giving rise to a coevolving network [70]. Endowing individuals with local prevalence-based awareness highlights the role of network effects [123],[122].

In this chapter, we first present SIS epidemic dynamics in its most basic form - as an ordinary differential equation describing a well-mixed population of infected and susceptible individuals (Section 3.1). The simplificity of the ODE formulation guides our intuition in a benchmark networked model of SIS dynamics, presented in Section 3.2 as a discrete-time Markov chain. Networked models of epidemic spreading more realistically captures the nuances of local spread, which is determined by the graph structure of the network. We then present our networked SIS process with dynamically distributed information and social distancing actions (Section 3.3). The information the agents receive comes from their social contacts and a global broadcast

about the current state of the epidemic. An agent's social distancing action reduces its contact network interactions, the magnitude of which depends on how informed it is. Using a mean-field approximation on infection probabilities, we prove that awareness reduces the endemic equilibrium level, but cannot improve upon the epidemic threshold for persistence (Section 3.4). In addition, we provide a stochastic comparison analysis between the awareness and benchmark processes using a monotone coupling technique (Section 3.5). This establishes the fact that adding awareness improves upon any epidemic performance metric in expectation compared to the benchmark model, e.g. time to eradication and the cumulative infected nodes over time. Through this analysis, we provide a closed-form expression for the difference in performance. Through extensive simulations on three families of random graphs (Erdős-Renyi, geometric, and scale-free), we find that local contact information most effectively limits epidemic spread (Section 3.6).

3.1 ODE models of SIS epidemics

The susceptible-infected-susceptible (SIS) model is described by the following two-state dynamics,

$$\begin{aligned} \frac{dS}{dt} &= \delta I - \frac{\beta SI}{N} \quad \text{Susceptible population} \\ \frac{dI}{dt} &= \frac{\beta SI}{N} - \delta I \quad \text{Infected population} \end{aligned} \tag{3.1}$$

where $\beta > 0$ is the infection rate of the disease, $\delta > 0$ is the healing rate, and $N > 0$ is the size of the population. The SIS model is one among many types of "compartmental" epidemic models, where each individual in the population is in a particular state, or compartment. Compartmental epidemic ODE equations of this type were first extensively studied by Kermack and McKendrick [51], where they studied the susceptible-infected-recovered (SIR) compartmental model. One can interpret (3.1) as having a continuum of agents of mass N, where the population is homogeneous and well-mixed. In other words, each individual interacts with every

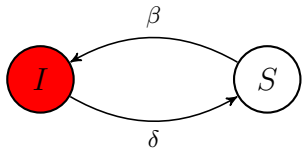


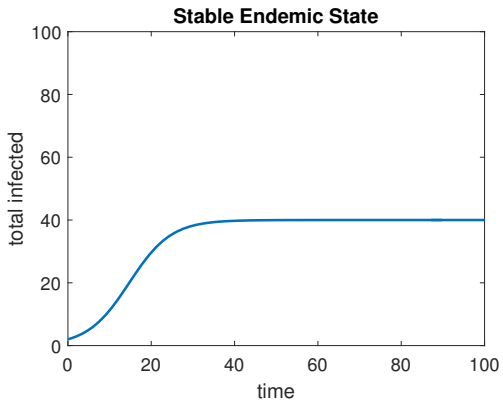
Figure 3.1: SIS dynamics state transition diagram

other individual in the population equally. Here, we assume that $S(0) + I(0) = N$ with $0 \leq S(0), I(0) \leq N$. The fact that $\frac{d}{dt}(S + I) = 0$ ensures $S(t) + I(t) = N$ with $S(t), I(t) \geq 0 \ \forall t$, so the total population is kept constant and the dynamics can be described by just one state variable. The infected population heals to become susceptible at a rate proportional to its level. The susceptible population becomes infected by interacting with the infected population (see Figure 3.1). In phase space, the point $(S^*, I^*) = (N, 0)$ (all-susceptible state) is always an equilibrium point - it is the all healthy state. The Jacobian matrix is

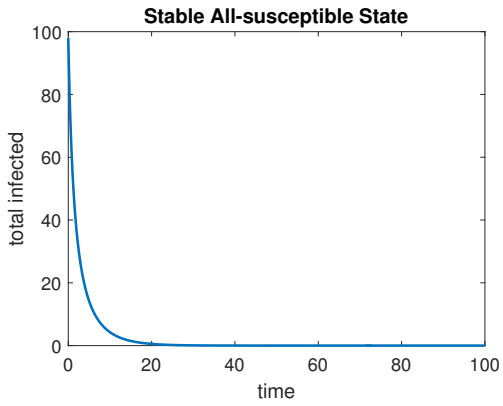
$$J = \begin{bmatrix} -\frac{\beta I}{N} & -\frac{\beta S}{N} + \delta \\ \frac{\beta I}{N} & \frac{\beta S}{N} - \delta \end{bmatrix}$$

Stability analysis shows it is stable if $R_0 \equiv \beta/\delta < 1$. Here, R_0 is known as the basic reproductive number in epidemiology, and quantifies the expected number of secondary infections from a single infected individual. Hence, $I(t) \to 0$. It is unstable if $R_0 > 1$. When (S^*, I^*) is unstable, there exists another equilibrium point $(\hat{S}, \hat{I}) = (\frac{\delta}{\beta}N, N(1 - \frac{\delta}{\beta}))$ which is stable. This equilibrium is known in epidemiology as the *endemic state*, where a fraction of the population is always infected. In this scenario, the disease is infectious enough to sustain itself within the population.

We already see that a simple mathematical model of epidemic spreading gives rise to a rich set of dynamics. Using standard techniques, the system (3.1) can be solved analytically.



(a) $I(t)$ when $\beta > \delta$. Here, $\beta = 0.5, \delta = 0.3, N = 100, I(0) = 2$.



(b) $I(t)$ when $\beta < \delta$. Here, $\beta = 0.5, \delta = 0.7, N = 100, I(0) = 98$.

Figure 3.2: Solution to SIS epidemic ODE. Left: the endemic equilibrium is asymptotically stable, so the disease persists. Right: the all-susceptible state is asymptotically stable, and the disease is eradicated.

3.1.1 Exact solution to the SIS ODE

One can derive the closed-form solution to (3.1). Since $S(t) = N - I(t)$, it amounts to solving a single variable ODE,

$$\frac{dI}{dt} = I\left(\beta(1 - \frac{I}{N}) - \delta\right).$$

We rearrange by dividing:

$$\frac{dI/dt}{I\left(\beta(1 - \frac{I}{N}) - \delta\right)} = 1$$

Let $c_N := N(\delta/\beta - 1)$. Using a partial fraction decomposition on the left-hand side, we obtain

$$\left(\frac{dI/dt}{I} - \frac{dI/dt}{I + c_N}\right) = -\frac{\beta}{N}c_N$$

Integrating from 0 to t with respect to the time variable, we get

$$[\log(I(t)) - \log(I(0)) - \log(I(t) + c_N) + \log(I(0) + c_N)] = -\frac{\beta}{N}c_N t$$

Taking the exponential,

$$\frac{I(t)}{I(0)} \times \frac{I(0) + c_N}{I(t) + c_N} = e^{(\beta - \delta)t}$$

Solving for $I(t)$, we obtain

$$I(t) = \hat{I} \times \frac{I(0)e^{(\beta-\delta)t}}{I(0)\left(e^{(\beta-\delta)t} - 1\right) + \hat{I}} \tag{3.2}$$

where we recall that $\hat{I} = N(1 - \frac{\delta}{\beta})$.

The solution (3.2) is consistent with the stability analysis of the previous section (see Figure 3.2). When $\beta < \delta$, the numerator vanishes for large t, so $I(t) \to 0$. When $\beta > \delta$, the exponential terms in the numerator and denominator dominate. Thus, $I(t) \to N(1 - \delta/\beta) = \hat{I}$. This holds for $0 < I(0) \leq N$.

3.2 A benchmark networked SIS model

Network models of epidemic spreading are more realistic than their well-mixed ODE counterparts, as they explicitly describe local transmission of disease. Here, a standard model of epidemic spreading over a finite static network of n agents is introduced (studied in [115],[3], and Section 5 of [2]), which we will refer to as the benchmark model.

Consider the set of nodes $\mathcal{N} = \{1, \ldots, n\}$ interconnected by a set of edges $\mathcal{E}$. Epidemic spread occurs in discrete time steps $t = 0, 1, \ldots$ over the undirected graph $\mathcal{G}_C = (\mathcal{N}, \mathcal{E})$, whose $n \times n$ adjacency matrix is defined for any $i, j \in \mathcal{N}$, as $[A_C]_{ij} = 1$ if $(i, j) \in \mathcal{E}$ and 0 otherwise. The graph $\mathcal{G}_C$ will be refered to as the contact network. An agent $i \in \mathcal{N}$ is either susceptible to the disease or infected by it. The epidemic states are defined as $\Omega = \{0, 1\}^n$. For any $s \in \Omega$ and $i \in \mathcal{N}$, either $s_i = 0$, meaning agent i is susceptible, or $s_i = 1$, meaning it is infected. Susceptible agents can contract the disease from neighboring agents in the contact network. When agent i is susceptible in the epidemic state is $s \in \Omega$ ($s_i = 0$), its probability of getting infected in the next time step due to an interaction with its neighbor $j \in \mathcal{N}_i^C$ is given by βs_j where $\beta \in (0, 1)$ is the transmission probability of the disease. Hence, an individual can only contract the disease from an infected neighbor. Agent i interacts with each of its neighbors independently. Therefore, i's probability of not becoming infected in

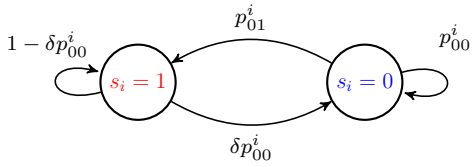


Figure 3.3: SIS dynamics state transition diagram

the next time step is

$$p^i_{00}(s) = \prod_{j \in \mathcal{N}_i^C} (1 - \beta s_j)$$

Consequently, its probability of getting infected is

$$p^i_{01}(s) = 1 - p^i_{00}(s) = 1 - \prod_{j \in \mathcal{N}_i^C} (1 - \beta s_j) \tag{3.3}$$

When agent i is infected in state $s \in \Omega$ ($s_i = 1$), it becomes susceptible in the next time step with probability $\delta p^i_{00}(s)$, where $\delta \in (0, 1)$ is the healing probability of the disease. Thus, for an infected node to become susceptible, it must heal and not get re-infected by its neighbors. Node i's transition probabilities are illustrated in Figure 3.3 and summarized by $\mathbb{P}_i : \Omega \times \{0, 1\} \to [0, 1]$ defined by

$$\begin{aligned} \mathbb{P}_i(s, 0) &= (1 - s_i) p^i_{00}(s) + s_i \delta p^i_{00}(s) \\ \mathbb{P}_i(s, 1) &= (1 - s_i) p^i_{01}(s) + s_i (1 - \delta p^i_{00}(s)) \end{aligned} \tag{3.4}$$

For each $i \in \mathcal{N}$ and $s \in \Omega$, the $\mathbb{P}_i$ define the benchmark SIS Markov chain over Ω by the $2^n \times 2^n$ transition matrix K with elements

$$K(s, s') \triangleq \prod_{i=1}^{n} \mathbb{P}_i(s, s'_i), \ \forall s, s' \in \Omega$$

This chain has one absorbing state, the all-susceptible state $\mathbf{o} \triangleq \{0\}^n$. The probability distribution of states $\pi^t \in \Delta(\Omega)$ at time $t = 0, 1, 2, \ldots$ (a stochastic row vector) evolves according to

$$\pi^{t+1} = \pi^t K$$

with initial distribution $\pi^0 \in \Delta(\Omega)$. Because there is one absorbing state (all susceptible), the stationary distribution is $\pi(\mathbf{o}) = 1$ which satisfies $\pi = \pi K$. The exact probability of node i being infected, $p_i(t)$ over time can be written

$$p_i^t = \sum_{s \in \Omega : s_i = 1} \pi^t(s) = \sum_{s \in \Omega : s_i = 1} (\pi^0 K^t)(s) \tag{3.5}$$

Consequently, the expected number of infected individuals at time t is

$$\sum_{i=1}^{n} p_i^t.$$

3.2.1 Mean-field approximation of the benchmark model

A mean-field approximation of this Markov chain is extensively studied in [1]. They consider the discrete-time deterministic dynamical system

$$\begin{aligned} x_i^{t+1} &= \psi_i(x) \\ &\equiv (1-\delta)x_i^t + (1-(1-\delta)x_i^t)\left(1 - \prod_{j \in \mathcal{N}_i^C} (1 - \beta x_j^t)\right), \ i \in \mathcal{N} \end{aligned} \tag{3.6}$$

Here, the state $x_i^t \in [0,1]$ approximates the true probability p_i^t (eqn. (3.5)) that node i is infected at time $t = 0, 1, 2, \ldots$. The advantage of using an MFA over analyzing the true infection probabilities is that we get a recursive update rule to determine the evolution of the approximations. The form of (3.6) comes from writing the expected value of the infection state s_i^{t+1} conditional on the previous state vector s^t. We represent the discrete-time mapping (RHS of (3.6)) with $\psi : [0,1]^n \to [0,1]^n$.

$$\begin{aligned} \mathbb{E}[s_i^{t+1} | s^t] = \mathbb{P}_i(s_i^t, 1) &= s_i^t(1 - \delta p_{00}^i(s^t)) + (1 - s_i^t) p_{01}^i(s^t) \\ &= s_i^t(1-\delta) + (1-(1-\delta)s_i^t) p_{01}^i(s^t) \\ &= s_i^t(1-\delta) + (1-(1-\delta)s_i^t)\left(1 - \prod_{j \in \mathcal{N}_i^C} (1 - \beta s_j^t)\right) \end{aligned}$$

Note that the above equation still depends on the true infection state s_i^t, and is not yet a mean-field approximation. The MFA (3.6) results by replacing the state s^t with x^t, thus defining a recursive update rule.

Remark 3.1. *A relation between the approximated probability of infection, x_i^t, and the true probability of infection, p_i^t, has yet to be established in the literature. The challenge is due to the product term $\left(1-\prod_{j\in\mathcal{N}_i^C}(1-\beta s_j^t)\right)$, which contains multiple cross terms of alternating sign in s_i and s_j, $j\in\mathcal{N}_i^C$. A mean-field approximation on p_i^t that does have an upper-bound guarantee follows the dynamics*

$$x_i^{t+1}=(1-\delta)x_i^t+(1-(1-\delta)x_i^t)\sum_{j\in\mathcal{N}_i^C}x_j^t,\ i\in\mathcal{N}. \tag{3.7}$$

These dynamics also fall into a class of models studied in [62, 1]. In fact, it is true that $\sum_{j\in\mathcal{N}_i^C}x_j^t\geq\left(1-\prod_{j\in\mathcal{N}_i^C}(1-\beta x_j^t)\right)$. The equation (3.7) still acts as a meaningful epidemic dynamic because the mapping is increasing in the x_j^t variables (infected neighbors increase my chance of getting infected).

The fact that x_i^t in (3.7) upper bounds p_i^t is proven by assuming independence of infection events (see Ch. 4 of [72]):

$$\begin{aligned}Pr(s_j^t=1,s_i^t=1)&=Pr(s_j^t=1|s_i^t=1)Pr(s_i^t=1)\\&\geq Pr(s_j^t=1)Pr(s_i^t=1)\\&=p_j^tp_i^t.\end{aligned}$$

3.2.2 Equilibrium properties of the benchmark SIS model

Here, we summarize a notable result on equilibrium properties of the benchmark MFAs (3.6) and (3.7) from [1]. All proofs in this subsection can be found in that reference, and assume connectedness of the contact network.

Theorem 3.1. *If $\lambda_{max}(\beta A_C+(1-\delta)I)<1$, the origin is globally asymptotically stable in the benchmark MFA.*

Here, $\lambda_{\max}(A)$ is the largest eigenvalue magnitude of a matrix A, A_C is the adjacency matrix of the contact network, and I is the $n\times n$ identity matrix. We can think of $\lambda_{\max}(\beta A_C+(1-\delta)I)$ as a networked version of R_0, the basic reproductive number.

Theorem 3.2 (Theorem 3.2 and 5.1 in [1]). *If* $\lambda_{max}(\beta A_C + (1-\delta)I) > 1$*, the discrete-time dynamical systems* (3.6) *and* (3.7) *each admit a unique equilibrium point* $q^* \in (0,1)^n$ *(non-zero* $\forall i$*). This equilibrium is asymptotically stable under the MFA* (3.6)*.*

Hence, just like in the ODE SIS model, there are two regimes of operation determined by an epidemic threshold - one in which the all-healthy state prevails, and one in which an endemic state prevails. The endemic state q^* is also referred to as a *non-trivial fixed point.*

3.3 The networked SIS model with awareness

We modify the benchmark Markov chain model of Section 3.2 to take into account the agent awareness of the current epidemic state. The information agent i receives comes from two sources: the proportion of infected neighbors in its local social network and a global broadcast of the proportion of infected nodes in the entire network. The social network is a graph $\mathcal{G}_I = (\mathcal{N}, \mathcal{E}_I)$ with the same nodes as $\mathcal{G}_C$ but with different edges, representing the nodes' social communication links. The set of i's neighbors in $\mathcal{G}_I$ is written $\mathcal{N}_i^I$. The information is given by

$$\mu_i(s) \equiv \frac{\alpha}{|\mathcal{N}_i^I|} \sum_{j \in \mathcal{N}_i^I} s_j + \frac{1-\alpha}{n} \sum_{j=1}^{n} s_j \ , \forall s \in \Omega \tag{3.8}$$

where $\alpha \in [0,1]$ is a parameter that governs the trust nodes place in information from their social contacts. Consequently, node i reduces its interactions with its physical neighbors through the social distancing action

$$a_i(s) \equiv 1 - \mu_i(s),$$

which reduces its susceptible-to-infected probability (3.3) to

$$p_{01,d}^i(s) \equiv 1 - \prod_{j \in \mathcal{N}_i^C} (1 - \beta a_i(s) s_j).$$

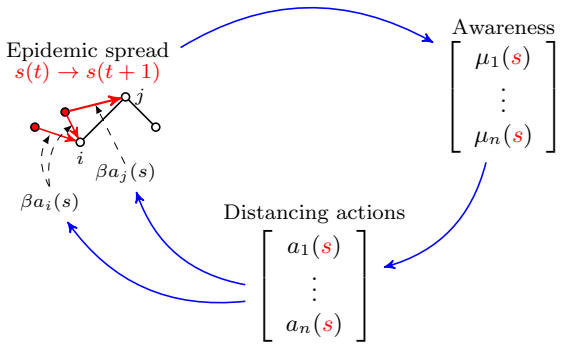


Figure 3.4: System-level diagram

We similarly define $p^i_{00,d}(s) \equiv 1 - p^i_{01,d}(s)$. An infected agent's probability of recovering becomes $\delta p^i_{00,d}(s)$. Note for all $s \in \Omega$, $p^i_{01,d}(s) \leq p^i_{01}(s)$. Combined with the social distancing behaviors a_i, the local awareness spread dynamics in (3.8) make the effect of user behavior on its infection probability endogenous to the benchmark chain model through a negative feedback loop. We define the $\mathbb{P}^d_i$ analogously to (3.4),

$$\mathbb{P}^d_i(s,0) = (1 - s_i)p^i_{00,d}(s) + s_i \delta p^i_{00,d}(s)$$
$$\mathbb{P}^d_i(s,1) = (1 - s_i)p^i_{01,d}(s) + s_i(1 - \delta p^i_{00,d}(s)).$$

Thus, $\mathbb{P}^d_i$ defines the distancing Markov chain over Ω by the transition matrix K_d with elements

$$K_d(s,s') \equiv \prod_{i=1}^{n} \mathbb{P}^d_i(s,s'_i), \ \forall s,s' \in \Omega$$

whose unique absorbing state is also $\mathbf{o} = \{0\}^n$, the all-susceptible state. The feedback loop between the epidemic state and agent awareness is illustrated in Figure 3.4.

The awareness model captures the different ways an agent may receive information about an ongoing epidemic from the media. Large media corporations and public health institutions such as the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO) often report an estimated total number of people infected nationwide or globally at a given time, and this information is disseminated amongst the population. Information is also exchanged through one's

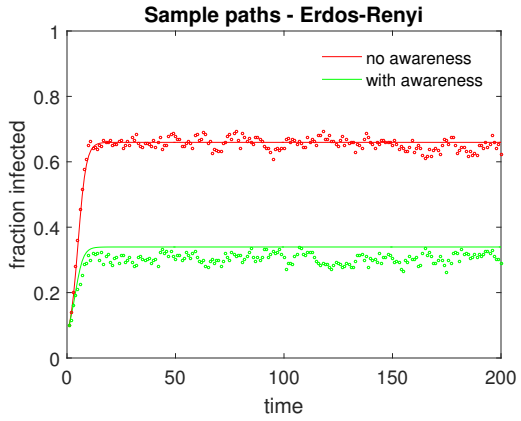


Figure 3.5: A pair of sample paths generated from the Markov chain models. Here, $\beta = 0.1, \delta = 0.5, \alpha = 1$, and $\mathcal{G}_I = \mathcal{G}_C$. Performed on a 100 node Erdős-Renyi network with parameter $p = 0.1$.

personalized social links, which can range beyond a person's geographic location.

It is important to note that since the all-susceptible state is the unique absorbing state for both benchmark and awareness chains. Since it is accessible from all other states, the chains must eventually absorb in finite time with probability one. This reveals the long-run equilibrium properties, but says nothing about what happens in the medium-run, before the disease dies out. The expected absorption time grows exponentially with the size of the network. Thus, one must wait an unrealistically long time for an epidemic to die out on its own.

3.4 The endemic state with awareness

In this section, we prove that an analogous MFA of the awareness epidemic model admits a non-trivial fixed point in the same regime of epidemic persistence of the benchmark model, $\lambda_{\max}(\beta A_C + (1-\delta)I) > 1$. Therefore, the awareness and social distancing mechanism cannot eradicate an epidemic. We prove existence of such an equilibrium, and it remains to prove that there is a unique non-trivial fixed point. Numerical simulations suggest the existence of a unique and asymptotically stable non-trivial fixed point in this regime.

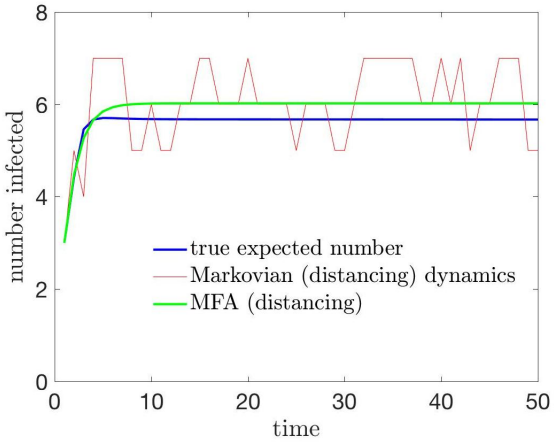


Figure 3.6: Comparison between the true expected number of infecteds, the approximated number of infecteds (MFA), and the exact Markovian dynamics. The MFA dynamics are calculated by computing the recursion $x_i^{t+1} = \phi_i(x)$ and summing over all nodes i. The true expected number of infecteds is calculated by iterating the probability distribution of states via $\pi^{t+1} = \pi^t K_d$, and summing over all states that give infected nodes. The MFA dynamics appear to overestimate the true expected number after a short amount of time. In this simulation, $\beta = 0.2$, $\delta = 0.1$, the contact network (as well as the information network) is the complete graph over $n = 8$ nodes, and $\lambda_{\max} = 7$. The initial probability distribution of states $\pi^0 \in \Delta(\Omega)$ as well as the initial state for the MFA and Markov chain is selected where three nodes are infected.

3.4.1 Mean-field approximation of the awareness model

In the same fashion in which (3.6) approximates the true probabilities of infection over time, we study a corresponding MFA of our awareness Markov chain. These dynamics will follow the update equation

$$\begin{aligned} x_i^{t+1} &= \phi_i(x) \\ &\equiv x_i^t(1-\delta) + (1-(1-\delta)x_i^t)\left(1 - \prod_{j \in \mathcal{N}_i^C} (1 - \beta a_i(s) s_j)\right) \end{aligned} \tag{3.9}$$

where again the state vector x^t is an approximation of the true infection probabilities p^t (3.5). Sample path trajectories of benchmark and awareness MFAs are shown in Figure 3.5. We have yet to establish an analytical relation between these quantities, see Remark 3.1 in Section 3.2.1. However, numerical simulations (Figure 3.6) suggest the MFA (3.9) is a slight overestimate of the true infection probabilities. Here, we

denote the awareness MFA mapping (RHS above) as $\phi : [0,1]^n \to [0,1]^n$.

We note a few properties about the MFAs mapping $\psi(x)$ (benchmark, eqn (3.6)) and $\phi(x)$ (awareness, eqn (3.9)). They are both nonlinear, continuous mappings satisfying $\phi(x) \prec \psi(x)$ for $x \in [0,1]^n \backslash \mathbf{0}_n$, $\phi(\mathbf{0}_n) = \psi(\mathbf{0}_n) = \mathbf{0}_n$. Linearization of ψ and ϕ about the origin yields the same Jacobian matrix, $\beta A_C + (1-\delta)I$, and hence the same linearized dynamics $x^{t+1} = (\beta A_C + (1-\delta)I)x^t$. The linear dynamics serve as an upper bound to both (3.6) and (3.9). Therefore, if $\lambda_{\max}(\beta A_C + (1-\delta)I) < 1$, the origin is a globally stable fixed point and it is an unstable fixed point if $\lambda_{\max}(\beta A_C + (1-\delta)I) > 1$.

3.4.2 Proof of existence of an endemic equilibrium

We now prove that the condition $\lambda_{\max}(\beta A_C + (1-\delta)I) > 1$ is sufficient for there to exist a non-trivial fixed point in the awareness MFA.

Theorem 3.1. *If $\lambda_{max}(\beta A_C + (1-\delta)I_n) > 1$, there exists a nontrivial fixed point p^* for ϕ. That is, $\phi(p^*) = p^*$ with $p^* \in (0,1)^n$.*

To re-iterate, the existence of such a fixed point suggests the epidemic has a metastable endemic state, where the spread of the disease is fast enough to sustain an epidemic in the network. Our condition coincides with the condition for existence, uniqueness, and global asymptotic stability of the non-trivial fixed point q^* of ψ, which is $\lambda_{\max}(\beta A_C + (1-\delta)I) > 1$, i.e. when the origin in the linearized dynamics is unstable (Theorem 5.1, [2]). This condition incorporates the factors that contribute to the rate of spreading - δ, β, and the contact network A_C. The proof of Theorem 3.1 makes use of the following three lemmas.

Lemma 3.1 (Lemma 3.1, [2])**.** *There exists a vector $\nu \succ 0_n$ such that $(\beta A_C - \delta I_n)\nu \succ 0_n$ if and only if $\lambda_{max}(\beta A_C + (1-\delta)I_n) > 1$.*

The connectedness assumption for $\mathcal{G}_C$ is necessary for the above Lemma because the proof applies the Perron-Frobenius theorem for nonnegative irreducible matrices. The next result is an equivalent formulation of Brouwer's fixed point theorem.

Lemma 3.2. *(Theorem 4.2.3, [48]): Suppose $f_i : D_n \to \mathbb{R}, i = 1, \ldots, n$ are continuous mappings, where*

$$D_n = \{x \in \mathbb{R}^n : x_i \in [\ell_i, u_i],\ \forall i\}$$

for real numbers ℓ_i, u_i. We also define the set

$$D_{-i} = \{x_{-i} \in \mathbb{R}^{n-1} : x_j \in [\ell_j, u_j]\ \forall j \neq i\}$$

If for every i and for all $x_{-i} \in D_{-i}$,

$$f_i(x_1, \ldots, \ell_i, \ldots, x_n) = f_i(x_{-i}, \ell_i) \geq 0 \tag{3.10}$$

$$f_i(x_1, \ldots, u_i, \ldots, x_n) = f_i(x_{-i}, u_i) \leq 0, \tag{3.11}$$

then there exists a point $x^* \in D_n$ such that $f_i(x^*) = 0$, for all $i = 1, \ldots, n$.

The final lemma needed is a technical result for the mean-field mappings ϕ_i.

Lemma 3.3. *For each $i \in \mathcal{N}$, define the maps $f_i : [0,1]^n \to \mathbb{R}$*

$$\begin{aligned} f_i(x) &\triangleq \phi_i(x) - x_i \\ &= -\delta x_i + (1 - (1-\delta)x_i) p^i_{01,d}(x) \end{aligned}$$

Then for any $i \in \mathcal{N}$ and $x_{-i} \in [0,1]^{n-1}$, the function $f_i(x_{-i}, \cdot)$ has a unique root $c_i^*(x_{-i}) \in [0,1)$ which depends continuously on x_{-i}. Furthermore, one can find a sequence $x^k_{-i} \to \mathbf{0}_{n-1}$ s.t. $c_i^*(x^k_{-i})$ is monotonically decreasing to 0.

Proof. For any $i \in \mathcal{N}$ and $x_{-i} \in [0,1]^{n-1}$,

$$f_i(x_{-i}, 0) = p^i_{01,d}(x_{-i}, 0) \geq 0.$$

and

$$f_i(x_{-i}, 1) = \delta(p^i_{01,d}(x_{-i}, 1) - 1) < 0.$$

The function $f_i(x_{-i}, \cdot)$ is strictly decreasing: for $a, b \in [0,1]$ s.t. $a < b$, $f_i(x_{-i}, a) - f_i(x_{-i}, b)$ is given by

$$\begin{aligned}(b-a)(\delta + (1-\delta)p^i_{01,d}(x_{-i}, b)) + \cdots \\ + (1 - a(1-\delta))(p^i_{01,d}(x_{-i}, a) - p^i_{01,d}(x_{-i}, b)) \\ > 0.\end{aligned}$$

This follows because $p^i_{01,d}(x_{-i}, x_i)$ is decreasing in x_i (x_i contributes to global awareness). Hence for every $x_{-i} \in [0,1]^{n-1}$, there is a unique $c_i^*(x_{-i}) \in [0,1)$ s.t. $f_i(x_{-i}, c_i^*(x_{-i})) = 0$, and $c_i^*(x_{-i})$ depends continuously on x_{-i}. To see this, observe that $c_i^*(x_{-i}) \in [0,1)$ is a root of the equation

$$(1 - (1-\delta)x_i)\left(1 - \prod_{j \in \mathcal{N}_i^C} (1 - a_i(x_{-i}, x_i)\beta x_j)\right) - \delta x_i = 0,$$

which is a polynomial in x_i. The coefficients of the polynomial depend continuously on $x_{-i} \in [0,1]^{n-1}$, and the roots of any polynomial are continuous with respect to its coefficients. Consequently, for any sequence $x^k_{-i} \to \mathbf{0}_{n-1}$, $c_i^*(x^k_{-i}) \to c_i^*(\mathbf{0}_{n-1}) = 0$ by continuity. This allows us to select a subsequence of x^k_{-i} such that c_i^* is monotonically decreasing along the subsequence. ■

We are now ready to prove the main result of this section.

Proof of Theorem 3.1: See Figure 3.7 for an illustration of the proof. Let the mappings f_i, $i \in \mathcal{N}$ be as in Lemma 3.3. We need to verify (3.10) and (3.11) hold for all i and for choices of ℓ_i, u_i satisfying $0 < \ell_i < u_i$. This will ensure the awareness dynamic ϕ has a fixed point other than the origin. For every $i \in \mathcal{N}$, choose $u_i = u$ where u satisfies

$$\max_{i \in \mathcal{N}} \max_{x_{-i} \in [0,1]^{n-1}} c_i^*(x_{-i}) < u < 1.$$

Then for all $x_{-i} \in [0,u]^{n-1}$,

$$f_i(x_{-i}, u) < 0$$

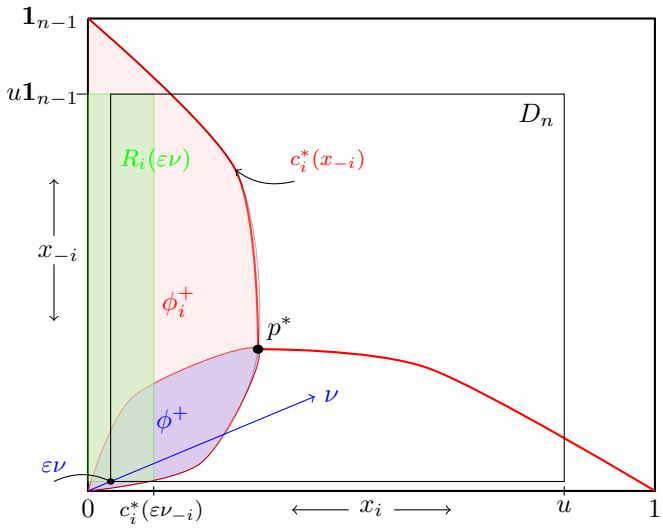


Figure 3.7: Diagram of the proof of Theorem 3.1. Here, p^* denotes a nontrivial fixed point of the awareness MFA mapping ϕ.

since $u > c_i^*(x_{-i})$. Thus, (3.11) holds, regardless of the choice of ℓ_i. It remains to find the $\ell_i > 0$ s.t. (3.10) is satisfied. Let $f(x) \triangleq [f_1(x), \ldots, f_n(x)]^T$ and define the sets

$$\phi_i^+ \triangleq \{x \in [0,1]^n : f_i(x) \geq 0\}, \quad \phi^+ \triangleq \bigcap_{j=1}^{n} \phi_j^+ \tag{3.12}$$

The Jacobian of f about the origin is $(\beta A_C - \delta I_n)$. By Lemma 3.1, there exists a vector $\nu \succ \mathbf{0}_n$ such that $(\beta A_C - \delta I_n)\nu \succ \mathbf{0}_n$. Consequently for sufficiently small $\varepsilon > 0$,

$$f(\varepsilon\nu) \succ \mathbf{0}_n,$$

or $\varepsilon\nu \in \phi^+$. We also define the set

$$R_i(x_{-i}) \triangleq \{y \in \mathbb{R}^n : y_i \in [0, c_i^*(x_{-i})], y_j \in [x_j, u], j \neq i\}$$

If $\varepsilon\nu_i \leq c_i^*(\varepsilon\nu_{-i})$ and $R_i(\varepsilon\nu) \subset \phi_i^+$, then (3.10) is satisfied, i.e. $f_i(x_{-i}, \varepsilon\nu_i) \geq 0$ on $D_{-i} = \{\varepsilon\nu_{-i} \preceq x_{-i} \preceq u\mathbf{1}_{n-1}\}$ for ε sufficiently small. We already have $\varepsilon\nu_i \leq c_i^*(\varepsilon\nu_{-i})$ because $\varepsilon\nu \in \phi^+ \subset \phi_i^+$. To show $R_i(\varepsilon\nu) \subset \phi_i^+$, by Lemma 3.3 we can find a sequence $\varepsilon_k\nu \in \phi^+$ with $\varepsilon_k \to 0$ s.t. $c_i^*(\varepsilon_k\nu_{-i})$ is monotonically decreasing to 0. By stopping at a large enough k, we can take a ε small enough such that

$$c_i^*(\varepsilon\nu_{-i}) = \min_{x_{-i} \in D_{-i}} c_i^*(x_{-i}) \tag{3.13}$$

Consequently, $R_i(\varepsilon\nu) \subset \phi_i^+$. Choosing ε small enough to satisfy (3.13) for all $i \in \mathcal{N}$ verifies (3.10) by using $D_n = \{x \in \mathbb{R}^n : x_j \in [\varepsilon\nu_j, u]\}$. By Lemma 2, ϕ has a fixed point contained in D_n. ■

The condition of Theorem 3.1 is independent of the awareness parameter α and the structure of the information network $\mathcal{G}_I$. Hence, social distancing alone cannot restore stability of the disease-free equilibrium point in the mean-field setting. This result is in contrast to conclusions made in different SIS model formulations [122, 102], where the epidemic threshold for stability of the disease-free state improves due to awareness. However, social distancing in our model lowers the overall metastable state of an epidemic.

Corollary 3.1. *If $q^* \succ \mathbf{0}_n$ is the unique nontrivial fixed point of ψ, then a nontrivial fixed point p^* of ϕ satisfies $p^* \prec q^*$ whenever $\alpha \in [0, 1)$.*

Proof. Define the sets ψ_i^+ and ψ^+ similarly as in (3.12) for the benchmark MFA. It was shown in [2] that q^* is the unique maximal element of ψ^+, i.e $q^* \succ q$, $\forall q \in \psi^+$, $q \neq q^*$. Observe $\phi(x) \prec \psi(x)$ for any $x \in [0, 1]^n \backslash \mathbf{0}_n$. Let $x \in \phi^+$, $x \neq \mathbf{0}_n$. Then $x \preceq \phi(x) \prec \psi(x)$, so $x \in \psi^+$. Therefore, $\phi^+ \subset \psi^+$, and since $p^* \in \phi^+$ for any nontrivial fixed point of ϕ, $p^* \prec q^*$. ■

The mean-field analysis of this section reveals the qualitative dynamics not addressed by an absorbing Markov chain analysis. Since the all-susceptible state is the unique absorbing state and accessible from every other state, the disease eradicates in finite time with probability one. This answers what happens in the long-run, whereas the MFA analysis answers what happens before this eventuality. The MFA analysis concurs with what is observed in simulations of the Markov chain dynamics - fast convergence to an endemic "metastable" state that persists for a very long time before eradication.

It remains to prove uniqueness and stability of the non-trivial fixed point guaranteed by Theorem 3.1. However, numerical simulations suggest stability of a unique nontrivial fixed point p^* of ϕ. A characterization of these fixed points for different random graph families is given in Section 3.6.

3.5 Stochastic comparison between benchmark and awareness models

This section presents a stochastic comparison analysis between the benchmark and awareness Markov chain models via a monotone coupling technique. This analysis shows that awareness reduces the expectation of any epidemic cost metric (e.g. time to extinction, total infected over time, etc) and the reduction can be given as a closed-form expression. First, relevant definitions and examples on monotone couplings are given. For a full reference on monotone coupling, see Ch 4 of [56].

3.5.1 Monotone couplings

Consider a general countable space X. A partially ordered set $(X, \preceq_X)$ is the set X together with a relation $\preceq_X$ among its elements which satisfies for all x, y, and $z \in X$,

- $x \preceq_X x$
- If $x \preceq_X y$ and $y \preceq_X z$, then $x \preceq_X z$
- If $x \preceq_X y$ and $y \preceq_X x$, then $x = y$.

Definition 3.1. *Let p_1, p_2 be probability measures on a measurable space $(X, \mathcal{F})$ and suppose $(X, \preceq_X)$ is a partially ordered set. A* monotone coupling *of p_1, p_2 is a probability measure p on $(X^2, \mathcal{F}^2)$ such that for all $x', y' \in X$,*

$$\sum_{x \preceq_X y'} p(x, y') = p_2(y') \text{ and } \sum_{y \succeq_X x'} p(x', y) = p_1(x'). \tag{3.14}$$

Thus, for any $x, y \in X$ s.t $x \not\preceq_X y$, $p(x, y) = 0$. A monotone coupling p can be thought of as a joint distribution on two possibly independent random variables whose

distributions are given by p_1, p_2, where one of them "dominates" the other in terms of a partial ordering.

Example 3.1. *Consider two biased coins A and B where the biases $q_A, q_B < 1$ for landing heads satisfy $q_A < q_B$. The independent distributions are given by $p_X(0) = 1 - q_X, p_X(1) = q_X$ for $X \in \{A, B\}$. The coupling is a joint distribution assigned to the pair of coin flips to ensure the q_A coin can never land heads with the q_B coin landing tails, while the marginal coin flip probabilities remain the same. Also, define $p_{AB} : \{0,1\}^2 \to [0,1]$ by*

$$\begin{cases} p_{AB}(0,0) &= 1 - q_B \\ p_{AB}(0,1) &= q_B - q_A \\ p_{AB}(1,0) &= 0 \\ p_{AB}(1,1) &= q_A \end{cases} \tag{3.15}$$

Checking (3.14)*, the marginals are such that $\sum_{b \geq 1} p_{AB}(1,b) = q_A$, $\sum_{a \leq 1} p_{AB}(a,1) = q_B$ and $\sum_{b \geq 0} p_{AB}(0,b) = 1 - q_A$, $\sum_{a \leq 0} p_{AB}(a,0) = 1 - q_B$. Thus, p_{AB} is a monotone coupling of p_A, p_B.*

A function $Z : X \to \mathbb{R}$ is *increasing in* X if whenever $x \preceq_X y$, $Z(x) \leq Z(y)$. The next result characterizes the difference in expectations of increasing random variables between the marginals of a monotone coupling.

Proposition 3.1. *Keeping the notation of Definition 3.1, suppose p is a monotone coupling of p_1, p_2. If a random variable $Z : X \to \mathbb{Z}_+$ is increasing in X, then*

$$\mathbb{E}_{p_2}(Z) - \mathbb{E}_{p_1}(Z) = \sum_{\tau=0}^{\infty} p(Z_\tau^c, Z_\tau) \tag{3.16}$$

where $Z_\tau = \{x : Z(x) > \tau\}$.

Proof. Consider the following quantities:

$$\begin{aligned} p(Z_\tau, Z_\tau) &= \sum_{x \in Z_\tau} \sum_{y \in Z_\tau} p(x,y) \\ &= \sum_{x \in Z_\tau} \sum_{y \succeq_X x} p(x,y) = p_1(Z_\tau) \qquad (3.17) \\ p(X, Z_\tau) &= p_2(Z_\tau) \end{aligned}$$

The second sum over $\{y \in Z_\tau\}$ can be replaced with $\{y \succeq_X x\}$ in (3.17) because 1) for any $x \in Z_\tau$, we have $\{y : y \succeq_X x\} \subset Z_\tau$; and 2) since p is a monotone coupling, for any $y \in Z_\tau$ s.t. $y \not\succeq_X x$, $p(x,y) = 0$. The last equality of (3.17) follows from (3.14). Since $(X, Z_\tau) \supset (Z_\tau, Z_\tau)$ we can write

$$\begin{aligned} p_2(Z_\tau) - p_1(Z_\tau) &= p(X, Z_\tau) - p(Z_\tau, Z_\tau) \\ &= p((X, Z_\tau) \backslash (Z_\tau, Z_\tau)) \\ &= p(Z_\tau^c, Z_\tau) \end{aligned}$$

Equation (3.16) immediately follows. ■

Example 3.2. *Consider the biased coins of Example 3.1. This example can be extended to sequences of $m \geq 2$ flips, $\{0,1\}^m$ with the partial order $x \preceq y$ if $x_i \leq y_i$, $i = 1, \ldots, m$, for $x, y \in \{0,1\}^m$. Define*

$$\begin{aligned} \overline{p}_X(x) &= \prod_{k=1}^{m} p_X(x_k), \;\; X \in \{A, B\} \\ \overline{p}_{AB}(x,y) &= \prod_{k=1}^{m} p_{AB}(x_k, y_k). \end{aligned}$$

Then $\overline{p}_{AB}$ is a monotone coupling of $\overline{p}_A, \overline{p}_B$. For $x \in \{0,1\}^m$, let $Z(x) = \sum_{i=1}^m x_i$ be the random variable of the number of heads for any given toss sequence. Then Z is increasing in $\{0,1\}^m$. By Proposition 3.1,

$$\mathbb{E}_{\overline{p}_B}(Z) - \mathbb{E}_{\overline{p}_A}(Z) = \sum_{\tau=0}^{m} \overline{p}_{AB}(Z_\tau^c, Z_\tau).$$

Of course, one could trivially compute the LHS above as $m(q_B - q_A)$ since the distribution of Z is Bernoulli. However, Proposition 3.1 generalizes the difference for any increasing $\mathbb{Z}_+$-valued random variable over a partially ordered set.

The notion of stochastic domination is a natural concept that arises in monotone couplings.

Definition 3.2. *An* upper set *$\mathcal{I}$ is a non-empty subset of $(X, \preceq_X)$ that satisfies the following property: if $x \in \mathcal{I}$ and $y \succeq_X x$, then $y \in \mathcal{I}$. Let p_1, p_2 be two probability measures on $(X, \mathcal{F})$. Then p_2* stochastically dominates *p_1, written as $p_2 \succeq p_1$, if for any upper set $\mathcal{I} \subset X$, $p_1(\mathcal{I}) \leq p_2(\mathcal{I})$.*

The comparison between benchmark and distancing chains falls into the framework of the above analysis.

3.5.2 Comparison over sample paths

Our main result provides a construction of a monotone coupling between the benchmark and distancing probability distributions on sample paths.

Definition 3.3. *A* sample path *is a sequence $g = \{g^t\}_{t \in \mathbb{Z}_+}$ such that $g^t \in \Omega$ and $K(g^t, g^{t+1}) > 0$ for all $t \geq 0$, and there is a $T < \infty$ such that $g^T = \mathbf{o}$. The set of sample paths is denoted by Γ.*

The *absorption time* $T : \Gamma \to \mathbb{Z}_+$ of a sample path g is given by

$$T(g) \triangleq \min\{t : g^t = \mathbf{o}\}. \tag{3.18}$$

Thus for all $g \in \Gamma$, $T(g) < \infty$. Also, $g^t = \mathbf{o}$ and $K(g^t, g^{t+1}) = 1$ for all $t \geq T(g)$. Note that Γ is countable since it is the countable union of the finite sets $\{g : T(g) = t\}$ for $t = 0, 1, 2, \ldots$.

The distribution $\mu_\pi : \mathcal{P}(\Gamma) \to [0,1]$ on sample paths under the benchmark SIS chain with starting distribution $\pi \in \Delta(\Omega)$ is given, for any $A \in \mathcal{P}(\Gamma)$, by

$$\mu_\pi(A) \triangleq \sum_{g \in A} \pi(g^0) \prod_{t=0}^{T(g)-1} K(g^t, g^{t+1})$$

and similarly under the distancing chain by

$$\nu_\pi(A) \triangleq \sum_{g \in A} \pi(g^0) \prod_{t=0}^{T(g)-1} K_d(g^t, g^{t+1}).$$

Also, note that Γ is defined to exclude the set of sample paths that are never absorbed, $\{g : g^t \neq \mathbf{o}, \forall t \in \mathbb{Z}_+\}$. These are infinite sequences that never terminate, and therefore are uncountable. The probabilities μ_π, ν_π, however are well-defined on Γ without such sample paths:

$$\begin{aligned}
\sum_{g \in \Gamma} \mu_\pi(g) &= \sum_{t=0}^{\infty} \mu_\pi(\{g : T(g) = t\}) \\
&= \sum_{s \in \Omega} \pi(s) \sum_{t=0}^{\infty} r_s(Q^t) - r_s(Q^{t+1}) \\
&= 1
\end{aligned}$$

Here, Q is the $2^n - 1 \times 2^n - 1$ sub-stochastic matrix of transition probabilities between non-absorbing states, and $r_s(Q)$ is the s^{th} row-sum of Q. Hence, $r_s(Q^t) - r_s(Q^{t+1})$ is the probability a sample path starting from state s is absorbed at time t. The elements of Q^t approach zero as $t \to \infty$.

Remark 3.2. *$(\Omega, \preceq_\Omega)$ is a partially ordered set. For $s, s' \in \Omega$, $s \preceq_\Omega s'$ if $s_i \leq s'_i$ for all $i \in \mathcal{N}$.*

Remark 3.3. *$(\Gamma, \preceq_\Gamma)$ is a partially ordered set. For $h, g \in \Gamma$, $h \preceq_\Gamma g$ if $h^t \preceq_\Omega g^t$ for all $t \in \mathbb{Z}_+$.*

Next, we present the main result, which constructs a monotone coupling distribution of ν_π, μ_π by exploiting the differences in node-level transition probabilities.

Theorem 3.1. *Suppose $x, y \in \Omega$ with $x \preceq_\Omega y$. For each $i \in \mathcal{N}$, define $\varphi_i^{x,y} : \{0,1\}^2 \to [0,1]$ according to*

$$x_i = y_i = 1, \quad \begin{cases} \varphi_i^{x,y}(0,0) & = \delta(1 - p_{01}^i(y)) \\ \varphi_i^{x,y}(0,1) & = \delta(p_{01}^i(y) - p_{01,d}^i(x)) \\ \varphi_i^{x,y}(1,0) & = 0 \\ \varphi_i^{x,y}(1,1) & = 1 - \delta(1 - p_{01,d}^i(x)) \end{cases} \tag{3.19}$$

$$x_i = y_i = 0, \quad \begin{cases} \varphi_i^{x,y}(0,0) & = 1 - p_{01}^i(y) \\ \varphi_i^{x,y}(0,1) & = p_{01}^i(y) - p_{01,d}^i(x) \\ \varphi_i^{x,y}(1,0) & = 0 \\ \varphi_i^{x,y}(1,1) & = p_{01,d}^i(x) \end{cases}$$

$$x_i = 0, y_i = 1, \begin{cases} \varphi_i^{x,y}(0,0) & = \delta(1 - p_{01}^i(y)) \\ \varphi_i^{x,y}(0,1) & = 1 - p_{01,d}^i(x) - \delta(1 - p_{01}^i(y)) \\ \varphi_i^{x,y}(1,0) & = 0 \\ \varphi_i^{x,y}(1,1) & = p_{01,d}^i(x) \end{cases} \tag{3.20}$$

Also, define $\varphi^{x,y} : \Omega^2 \to [0,1]$ for $x \preceq_\Omega y$ by

$$\varphi^{x,y}(\omega, \omega') \triangleq \prod_{i=1}^{n} \varphi_i^{x,y}(\omega_i, \omega_i') \quad \forall \omega, \omega' \in \Omega.$$

Lastly, define $\Phi_\pi : \Gamma^2 \to [0,1]$ for any $\pi \in \Delta(\Omega)$ by

$$\Phi_\pi(h, g) \triangleq \chi(h^0 = g^0)\pi(h^0) \prod_{t=0}^{T(g)-1} \varphi^{h^t, g^t}(h^{t+1}, g^{t+1}).$$

Then Φ_π is a monotone coupling of ν_π, μ_π.

Proof. See Appendix. ■

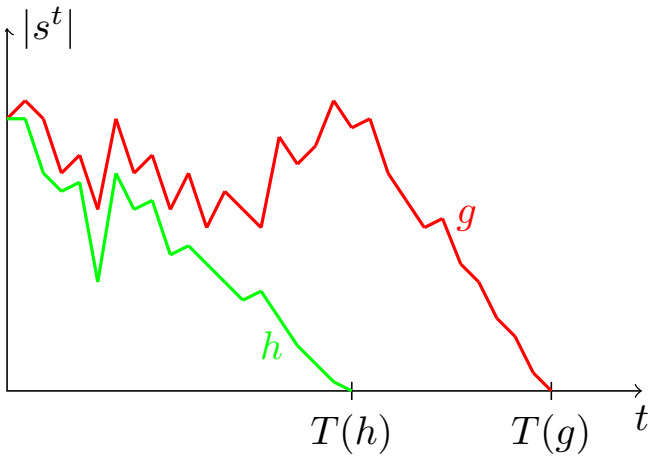


Figure 3.8: A pair of sample paths (h, g) drawn from Φ_π.

The coupling between node-level transition probabilities $\varphi_i^{x,y}$ given in (3.19)-(3.20) are used to establish a coupling between benchmark and distancing probability distributions on sample paths. The form of the $\varphi_i^{x,y}$ is identical to the method in which two biased coins are coupled in (3.15). When the coupling rule is applied to each node's probability of infection, it ensures no node can be infected in the distancing chain while being susceptible in the benchmark chain. Consequently, the monotone coupling Φ_π is a distribution on pairs of sample paths (h, g) satisfying $h^0 = g^0$, $h \preceq_\Gamma g$ (see Figure 3.8) and marginally, $h \sim \nu_\pi$, $g \sim \mu_\pi$. The next result characterizes the difference between μ_π and ν_π-expectations of any non-negative increasing function on the sample paths Γ with respect to the coupling distribution Φ_π.

Corollary 3.2. *For any increasing $\mathbb{Z}_+$-valued random variable Z in Γ,*

$$\mathbb{E}_{\mu_\pi}(Z) - \mathbb{E}_{\nu_\pi}(Z) = \sum_{\tau=0}^{\infty} \Phi_\pi(Z_\tau^c, Z_\tau). \tag{3.21}$$

where $Z_\tau = \{x \in \Gamma : Z(x) > \tau\}$.

Proof. Immediate from Theorem 3.1 and Proposition 3.1. ■

One can think of an increasing $Z : \Gamma \to \mathbb{Z}_+$ as an epidemic cost metric. Here, $\Phi_\pi(Z_\tau^c, Z_\tau)$ is simply the probability a benchmark sample path g incurs a cost greater

than τ, while the corresponding distancing sample path h costs less than τ, where $(h, g) \sim \Phi_\pi$. Some examples of increasing $\mathbb{Z}_+$-valued random variables in Γ are

- The absorption time $T : \Gamma \to \mathbb{Z}_+$, defined by (3.18).
- Cumulative infected individuals up to time m, defined by $g \mapsto \sum_{t=0}^{m} |g^t|$, where $|s| \triangleq \sum_{i \in \mathcal{N}} s_i$ for $s \in \Omega$.
- The "epidemic spread", or how many unique nodes that contract the disease in a given amount of time m. This is given by $g \mapsto \sum_{i \in \mathcal{N}} \chi_{E_i}(g)$, where $E_i = \{g : \sum_{t=0}^{m} g_i^t > 0\}$.

The difference (3.21) encodes many complex dependencies on the epidemic parameters δ and β, the awareness weight α, and the structure of the graphs $\mathcal{G}_C$ and $\mathcal{G}_I$.

The following result establishes stochastic domination of the distancing chain by the benchmark chain.

Corollary 3.3. *The benchmark chain stochastically dominates the distancing chain on sample paths, i.e.* $\mu_\pi \succeq \nu_\pi$.

Proof. For any upper set $\mathcal{I} \subset \Gamma$, $\chi_{\mathcal{I}}(\cdot)$ is increasing in Γ. By (3.21),

$$\begin{aligned} \mu_\pi(\mathcal{I}) - \nu_\pi(\mathcal{I}) &= \mathbb{E}_{\mu_\pi}(\chi_{\mathcal{I}}) - \mathbb{E}_{\nu_\pi}(\chi_{\mathcal{I}}) \\ &= \Phi_\pi(\mathcal{I}^c, \mathcal{I}) \geq 0 \end{aligned}$$

■

This result affirms the intuition that sample paths with consistently higher numbers of infected individuals occur with higher probability in the benchmark chain than in the distancing chain.

3.6 The value of contact information

We illustrate through numerical simulations how the structure of the contact network $\mathcal{G}_C$ affects the course of an epidemic in the presence of awareness and social distancing. Extensive analytical and simulation studies have been conducted without awareness [115, 36, 30, 109, 81]. Here, we look at three random graph families - geometric, Erdős-Rényi, and scale-free. These networks are relevant in studying epidemic spreading because they exhibit a variety of qualitative features that reflect real-world networks. Geometric networks portray people connected by geographic distance. Erdős-Rényi random networks display a small-world effect common in many real world networks - e.g neural and social influence networks. Online social networks and the World Wide Web are examples of scale-free networks [69].

In our model, the social network $\mathcal{G}_I$ is generated directly from $\mathcal{G}_C$ via a parameter $p \in (0, 1)$ through the following procedure: 1) Select a fraction p of existing edges in $\mathcal{E}_C$ at random and remove them from the edge set; 2) For each of the selected edges, select one of the two end nodes randomly (e.g with probability 1/2) as the root node; 3) For each of the selected root nodes i, select $j \neq i$ uniformly at random and add the edge (i, j). For p close to one, the resulting graph $\mathcal{G}_I = (\mathcal{N}, \mathcal{E}_I)$ exhibits the small-world effect (small average shortest path length and small clustering) [117]. When $p = 0$, $\mathcal{G}_I = \mathcal{G}_C$.

For the contact networks, geometric random graphs are generated by placing n points uniformly at random on the unit torus (unit square with periodic boundary conditions). An edge exists between any two points if they are less than a specified distance $r \in (0, 1)$ away. Erdős-Rényi random graphs are constructed by forming an edge between any two nodes independently with a fixed probability $p_{\text{ER}} \in (0, 1)$. Scale-free networks are generated by the preferential attachment algorithm [12]: starting with an initial connected graph of $m_0 \geq m$ nodes, $n - m_0$ additional nodes are added sequentially with each incoming node establishing links to m existing nodes in

the network. The probability a node receives an incoming link is proportional to its degree. We performed simulation analysis on one network from each random graph family. The networks all have 1000 nodes with an average degree of 10, and hence the same number of edges. This allows us to isolate the effect different network structures have on epidemic spreading.

In Figure 3.9, the normalized non-trivial fixed points are characterized for the three random networks in the interval of epidemic persistence. The norms of these points indicate the size of the endemic states and they slightly overestimate the actual long-run infected fraction observed in stochastic simulations of the Markov chains. The condition of Theorem 3.1 guarantees persistence of the epidemic in both benchmark and awareness models in the mean-field setting. However, the first-order condition of Theorem 3.1 serves as an upper bound of the exact Markovian epidemic threshold with respect to the ratio δ/β. Furthermore, the exact threshold is observed to increase with the addition of awareness.

In Figure 3.10, we quantify the "epidemic spread" by the number of unique nodes that contract an infection as time progresses when one uniformly random node is initially infected. This metric is an example of an increasing random variable over sample paths (Section 3.5) and is helpful in revealing not only how fast an epidemic initially spreads in the network, but also how far-reaching it is. A key observation is that contact awareness ($p = 0, \alpha = 1$) slows epidemic spreading better than any other awareness configuration at the beginning of an epidemic. This is intuitively clear since contact awareness provides nodes with the most vital information if they are in danger of getting infected. As p increases, $\mathcal{G}_I$ deviates more from $\mathcal{G}_C$ and the information nodes receive become less vital.

Erdős-Rényi and scale-free (with $m = 5$) networks admit disease spread throughout the entire network in a short amount of time, even with social distancing (Figure 3.10a, 3.10c). This is attributed to small average shortest path lengths (Ch. 8 &

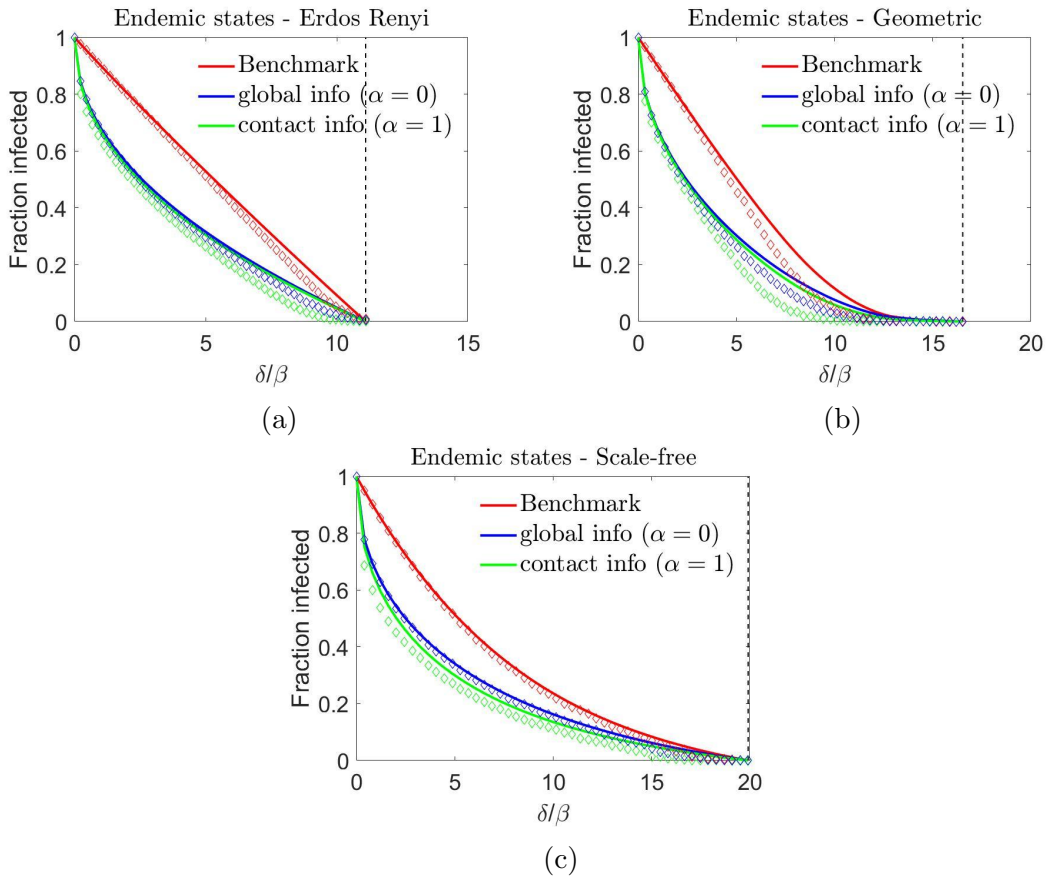


Figure 3.9: Norms of the nontrivial fixed points (solid lines) and long-run fraction of infected in stochastic simulations (diamonds) in the range of epidemic persistence, $\delta/\beta \in [0, \lambda_{\max}(A_C)]$, for $n = 1000$ node networks. The fixed points are computed by iterating the MFA dynamics ψ and ϕ starting from the all infected state (all-ones n-vector). The stochastic long-run infected fractions are computed by averaging the levels of epidemic states in the latter half of a sample run of length 200, starting from the all-infected state. Vertical dashed lines indicate $\lambda_{\max}(A_C)$. (a) Erdős-Rényi random network with $p_{\text{ER}} = .01$, $\lambda_{\max}(A_C) = 11.1$. Here, $p_{\text{ER}} > \log n/n$, the regime where the network is connected with high probability. (b) Geometric random graph with $r = .0564$, $\lambda_{\max}(A_C) = 16.52$. (c) Scale-free generated from the PA algorithm with $m = 5$, $\lambda_{\max}(A_C) = 19.9$. The parameters are chosen such that all networks have the same average degree $d = 10$.

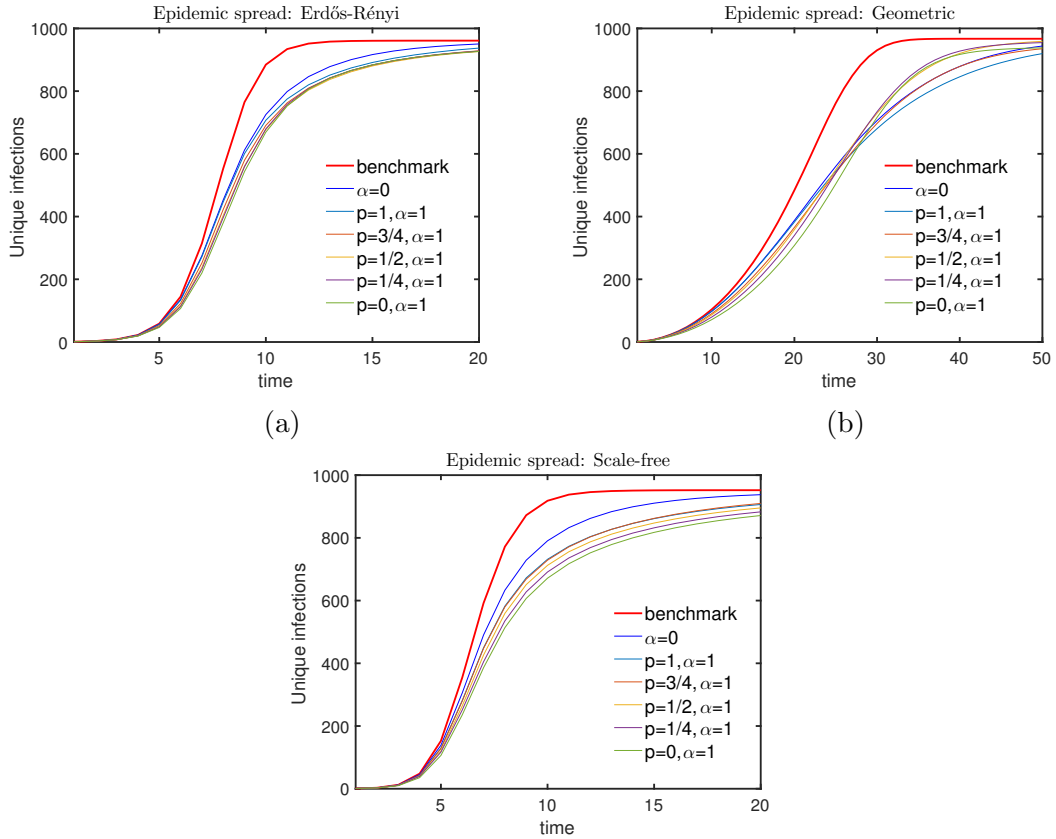


Figure 3.10: Epidemic spreading as a function of time (same networks and simulation setup as Fig. 3.9). Local contact information (α near 1, p near 0) slows spread most effectively for (a),(c), and the early stages of (b), whereas global information (α near 0 or $\alpha = 1, p$ near 1) is least effective. Note the inversion of awareness effectiveness in (b). In these simulations, $\delta = \beta = 0.2$.

12, [69]), allowing the epidemic to quickly spread to other parts of the network. Random geometric networks are characterized by high clustering and large diameter. Clustering slows the spread of an epidemic (Figure 3.10b), but also contributes to increasing the final epidemic size [112]. The virus stays localized and spreads slowly. This explains the inversion of awareness parameters in Figure 3.10b. By the time the epidemic first reaches its endemic level around $t = 20$, many nodes have not yet been exposed because at this point their local communities are untouched. Thus, having global or long-range social awareness (low α or high p) is more beneficial over contact awareness.

3.7 Discussion

In this chapter, we modified the benchmark networked SIS epidemic process to include agent awareness, where prevalence-based information comes from social contacts and a global broadcast of the overall infected fraction. Agents take social distancing actions based on the level of information received, which reduces their probabilities of getting infected. We showed that awareness does not change the epidemic threshold for persistence by proving existence of a nontrivial fixed point in the mean-field approximation. Any nontrivial fixed point of the distancing model is strictly component-wise less than the unique nontrivial fixed point of the benchmark model. Therefore, awareness reduces the endemic infection level, but it cannot eradicate the epidemic. We provided a full stochastic comparison analysis between the benchmark and distancing chains in terms of their respective probability distributions on sample paths by constructing a monotone coupling. Consequently, adding awareness reduces the expectation of any increasing random variable on sample paths and we obtain a closed form expression for the reduction. This analysis shows that awareness reduces the expectation of any epidemic cost metric (e.g. time to extinction, total infected

over time, etc) and the reduction can be given as a closed-form expression. In simulations, we showed epidemic spreading depends on the network structure. In particular, qualitative features such as small-world effects, clustering, and diameter explain the results seen in simulations. We also found that local contact awareness is the most effective at slowing epidemic spread, and global awareness is the least effective.

The importance of local contact information is highlighted through our simulations, and also corroborates with similar studies of the effect of awareness on epidemic spread [85, 103]. It is unlikely that an individual would know the infection status of those he comes into physical contact with. Attaining such information would require great effort on the part of centralized health organizations such as the World Health Organization and the Center for Disease Control. In particular, one method that attempts to gather such information is known as contact tracing. This involves identifying the diagnosis of the people an infected individual comes into contact with. This process can be quite challenging, requiring a full scale effort for highly contagious diseases like Ebola [6, 18].

The reactive social distancing mechanism in our model is a heuristic approximation of individual decision-making that emphasizes their risk-averseness, but does not take into account other factors such as financial burdens and the actions of their neighbors. For example, if someone knows his neighbor is sick and is staying home, there is no reason for him not go out to work that day. Similarly, if an individual is sick, he will likely stay at home to recover and to not risk spreading the sickness in public. These tradeoffs are studied in [31] and [92], in which game interactions and disease dynamics are coupled. Furthermore, our model does not consider higher order decision-making, such as non-linear distancing functions or basing actions on memory of past infections. For instance, the intensity of activity reduction can improve upon the epidemic threshold [94, 123]. Additionally, knowledge of epidemic severity in the past can influence the coupling of disease and vaccination behavior

[21], giving rise to oscillatory dynamics. In this vein, an area of future work would be to establish general classes of awareness-induced behavior mechanisms and their resulting asymptotic outcomes.

The choice to vaccinate is the focus of other models of decision-making in epidemic networks [21, 65, 68]. Here, studying the effect of social distancing in isolation from other preventative measures like vaccination helps to inform the limits of its effectiveness in reducing epidemic spread [92]. This knowledge is especially valuable in situations where social distancing actions are relatively cheap (e.g. washing hands more frequently) compared to producing and distributing vaccines. Furthermore, social distancing can act as an effective control measure in the absence of a developed vaccine.

CHAPTER IV

INFLUENCE OF OPINIONS FOR ENVIRONMENTAL CONSERVATION

A tragedy of the commons occurs when individuals in a population are driven by their own selfish interests, resulting in the depletion of a common resource on which they all depend. The interactions that drive such tragedies are modeled in classical game theory as a prisoner's dilemma [25, 73, 119]. The rational choice for an individual is to defect, regardless of what others are doing. However, classical models do not account for the consequences of action - individual actions affect the environment. Consequently, the state of the environment may shape individual incentives for future action. Dynamical models of these coevolutionary features have been developed to understand general conditions under which tragedies will occur or be averted [119, 17, 95]. Similarly, the study of common-pool resource games suggest that rational play among larger populations leads to resource collapse with higher probability [43, 91].

In his landmark paper [40], Hardin argues that such tragedies are inevitable given a growing human population, unless preventative measures are taken. To address the problem of preventing tragedies, there has been speculation about what intervention strategies will be effective. Interventions from centralized government entities are called for, through implementing and enforcing new policies restricting overconsumption [73, 8]. For example, imposing taxes on resource usage may provide a financial deterrent to overuse [44]. Passing regulatory laws on fishers gives fish populations a chance to recover [46]. Hence, such direct intervention policies provide the incentives necessary to instigate conservation behaviors [82].

Information also plays an important role. Individuals may not take pro-environmental

actions if they are not informed about why such actions are necessary [113]. Environmental awareness and education can lead to behavior changes when individuals realize that environmental degradation has adverse effects on their own community or household. For example, information from household metering about the severity of water scarcity drove efforts to conserve water [110]. However, statistics and facts may be ineffective to instigate behavior changes if such issues are politicized [58]. In these situations, public opinion is susceptible to propaganda from news outlets and social media. Environmental information is necessary to affect behavior change, but may not be sufficient [106, 82, 113]. The efficacy of these proposed solutions are rarely tested using dynamical models that couple actions and environmental changes [73].

A taxation mechanism on resource investment was studied in the setting of a common-pool resource game where under certain conditions, higher tax rates can lead to lower probability of resource collapse [44]. However, asymptotic outcomes are not considered in this static one-shot game. In a recent work [59], an infinite horizon optimal control framework was applied to a dynamical model[1] to identify conditions under which an optimal prescribed consumption rate ensures resource sustainability. However, the consumption rate is not directly manipulated by taxing, pricing, or other social control policies.

In contrast, we consider in this chapter such direct control policies. We formulate optimal control problems that study the role of information and incentive-based intervention policies with the objective of maximally conserving the environmental state over a finite time horizon. We apply these control formulations to the model of ref. [119], due to its general framework. It models a population of myopic individuals whose actions affect and are affected by the environment. This framework

[1] Those dynamics can be reduced to a linear system by an appropriate transformation. This differs fundamentally from the dynamical system considered in this chapter, which is highly nonlinear and cannot be transformed into a linear system.

differs from that of differential games [13], where individuals select strategies to maximize long-term payoffs given action-dependent dynamic environments. To implement information-based control policies, we introduce a dynamic public opinion that imperfectly tracks the true environmental state. We present two formulations in which the control directly affects public opinion: propaganda strategies that perturb public opinion, and awareness-raising strategies where learning of the true environmental state is encouraged. Additionally, we formulate an incentive control problem by allowing an external entity to influence the population's incentive to cooperate together. In all three formulations, we compute optimal controls by numerical means (by "optimal" in this chapter, we mean locally optimal since the problems we formulate are nonconvex).

The main contributions and findings of this chapter are 1) the formulation of optimal control problems to address the tragedy of the commons through direct policy interventions and 2) the solutions of these problems, obtained by numerical techniques, result in highly oscillatory behavior. In particular, we show through simulations that the objectives of the formulated problems are achieved, at the expense of inducing highly variant dynamics characterized by oscillatory cycles between low and high resource states. First, we briefly review some preliminaries, and the details and main findings of the game-environment feedback model of [119].

4.1 The prisoner's dilemma

The rationality of defection or consumption is best explained quantitatively by the *prisoner's dilemma* (PD) game. It is a standard model utilized in game theory to show that following selfish interests result in sub-optimal outcomes for all. The PD

game is played between two players as described by the following payoff matrix.

$$\begin{array}{c} \\ \mathcal{C} \\ \mathcal{D} \end{array} \begin{array}{c} \begin{array}{cc} \mathcal{C} & \mathcal{D} \end{array} \\ \begin{bmatrix} R & S \\ T & P \end{bmatrix} \end{array} \tag{4.1}$$

Players choose between two actions, cooperate ($\mathcal{C}$) or defect ($\mathcal{D}$). When both players cooperate (strategy profile $(\mathcal{C}, \mathcal{C})$), both are rewarded a payoff R. When a cooperator plays against a defector $(\mathcal{C}, \mathcal{D})$, it receives the sucker's payoff S while the defector receives the tempting payoff T. When both players defect $(\mathcal{D}, \mathcal{D})$, they receive a punishment payoff P. The payoff structure follows $T > R > P > S$, and without loss of generality, we assume all payoff parameters are positive. A player's optimal action is $\mathcal{D}$ regardless of the other player's action. Hence, the only pure Nash equilibrium is mutual defection, $(\mathcal{D}, \mathcal{D})$. The rational choice induces a suboptimal outcome since they would receive a better payoff (R) if they had both cooperated.

4.2 Co-evolutionary game theory

In this section, we outline the formulation of the dynamical model of [119]. It modifies a replicator equation of two strategies (cooperate and defect) by introducing environmental feedback. The actions of individuals in the population have consequences on the environment and as a result, shape individual incentives in the future.

4.2.1 The replicator dynamics

The replicator equation describes evolutionary dynamics of strategy frequencies in a well-mixed, infinite population [118, 99]. Under this dynamic, strategies proliferate in the population if their fitnesses exceed the average population fitness, and decline if they do not. Hence, success in the replicator dynamic is relative.

In its most general form, the replicator dynamics for n strategies are given by

$$\dot{x}_i = x_i(f_i(x) - \bar{f}(x)), \ i = 1, \ldots, n \tag{4.2}$$

where $x_i \geq 0$ is the frequency of strategy i in the population, $f_i(x)$ is the expected fitness to an individual in the population adopting strategy i when the population state is $x = [x_1, \ldots, x_n]$, and $\bar{f}(x) = \sum_{i=1}^{n} x_i f_i(x)$ is the average fitness of the population. Since $\sum_i \dot{x}_i = 0$, the population size remains constant. Without loss of generality, $\sum_i x_i(t) = 1, \forall t$. Hence, the population dynamics evolve over the n-dimensional simplex: $x(t) \in \Delta_n \equiv \{x \in \mathbb{R}^n | \sum_i x_i = 1, x_i \geq 0\}, \forall t$. A strategy i is said to be *extinct* if $x_i = 0$. By (4.2), an extinct strategy cannot re-enter the population.

To illustrate the replicator equation for the prisoner's dilemma game (4.1), we note there are two strategies, cooperate ($\mathcal{C}$) or defect ($\mathcal{D}$). Therefore, we can describe the composition of the population with a single state x, the fraction of cooperators. The fraction of defectors is then $1 - x$. A proportion x of interactions of an individual are with cooperators, and $1-x$ with defectors. Consequently, the fitnesses of a cooperator and defector are

$$f_{\mathcal{C}}(x) = Rx + S(1 - x) \quad \text{(cooperator fitness)}$$

$$f_{\mathcal{D}}(x) = Tx + P(1 - x) \quad \text{(defector fitness)}$$

Using (4.2), the evolution of cooperators is

$$\begin{aligned} \dot{x} &= x(f_{\mathcal{C}}(x) - \bar{f}(x)) \\ &= x(f_{\mathcal{C}}(x) - x f_{\mathcal{C}}(x) - (1 - x) f_{\mathcal{D}}(x)) \\ &= x(1 - x)(f_{\mathcal{C}}(x) - f_{\mathcal{D}}(x)) \end{aligned}$$

Given the payoff structure $R < T$ and $S < P$, $f_{\mathcal{C}}(x) < f_{\mathcal{D}}(x)\ \forall x$, and hence $\dot{x} < 0$. Therefore, cooperation dies out in the population and defectors dominate.

4.2.2 A model of replicator dynamics with environmental feedback

The replicator equation provides evolutionary outcomes for the games often studied in classical game theory such as the prisoner's dilemma. It does not, however, address the long-term impact individual actions have on the environment. Defectors who consume public resources without cooperating to conserve it will eventually have no

resource to consume. Realistically, incentives for future actions change as a result of re-shaping the environment. Here, we introduce the model of [119], which incorporates a game-environment feedback into the replicator equation. This model is intended to provide a general framework in which to portray the dynamics of tragedy of the commons scenarios.

The model introduces an "environment" variable $n \in [0,1]$ where $n = 0$ means the environment is in a deplete state and $n = 1$ means it is in a replete state. The game-environment coupled dynamics obey the following differential equations.

$$\dot{x} = x(1-x)(f_{\mathcal{C}}(x,n) - f_{\mathcal{D}}(x,n))$$

$$\dot{n} = n(1-n)(-1 + (1+\theta)x)$$

Here, the $n(1-n)$ term indicates a logistic growth of the environmental state, and serves to constrain the dynamics $n(t) \in [0,1]$. The growth or decline of the environment depends on the frequency x of cooperators in the population, who restore n at a rate $\theta > 0$. We assume the payoff matrix A_n that underlies individual interactions within the population is a function of n, and takes the form

$$A_n = \begin{bmatrix} R_n & S_n \\ T_n & P_n \end{bmatrix} \equiv n \begin{bmatrix} R_1 & S_1 \\ T_1 & P_1 \end{bmatrix} + (1-n) \begin{bmatrix} R_0 & S_0 \\ T_0 & P_0 \end{bmatrix}.$$

When $n = 1$, we impose that defection is the dominant strategy, that is, $R_1 < T_1$ and $S_1 < P_1$. Hence the only pure Nash equilibrium in the game defined by the payoff matrix A_1 is mutual defection, where players obtain a payoff P_1. We leave the structure of the game in the depleted state, defined by the payoff matrix A_0, unconstrained in order to explore the possible asymptotic outcomes of the system. The coevolutionary process between strategy frequencies and environmental state is illustrated in Figure 4.1.

Thus, the fitnesses for cooperators and defectors are

$$f_{\mathcal{C}}(x,n) = R_n x + S_n(1-x) \quad \text{(cooperator fitness)}$$

$$f_{\mathcal{D}}(x,n) = T_n x + P_n(1-x) \quad \text{(defector fitness)}.$$

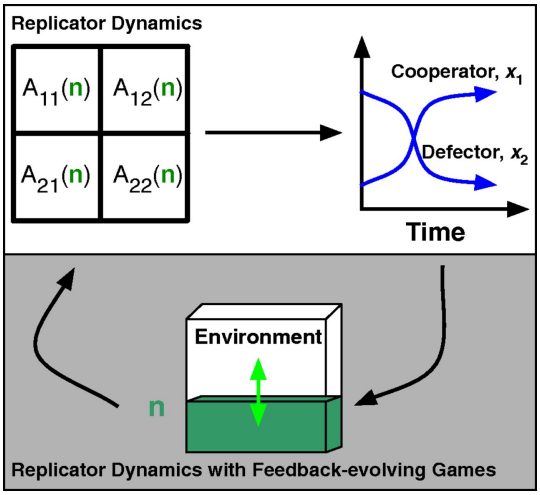


Figure 4.1: (Reproduced from [119]) Illustration of the feedback between strategy frequencies x and environmental state n. The payoff matrix is a function of n and determines the evolution of strategists in the population. As a result, the change in strategists influences the environmental state, causing a feedback loop of coevolutionary dynamics.

and the difference of the fitness of cooperators to defectors is

$$\begin{aligned} g(x,n) &\equiv f_{\mathcal{C}}(x,n) - f_{\mathcal{D}}(x,n) \\ &= xn((R_1 - S_1 - T_1 + P_1) - (R_0 - S_0 - T_0 + P_0)) + x(R_0 - S_0 - T_0 + P_0) \\ &\quad + n((P_0 - S_0) - (P_1 - S_1)) - (P_0 - S_0) \end{aligned} \tag{4.3}$$

We denote the state vector $\boldsymbol{y}(t) \equiv [x(t), n(t)]^\top$ and the system mapping as $F : [0,1]^2 \to \mathbb{R}^2$. In this notation, we re-write the system dynamics as

$$\begin{aligned} \dot{\boldsymbol{y}} \equiv F(\boldsymbol{y}) &= \begin{bmatrix} x(1-x)g(x,n) \\ n(1-n)(-1+(1+\theta)x) \end{bmatrix} \\ x(0) = x_0 \in [0,1], &\quad n(0) = n_0 \in [0,1] \end{aligned} \tag{4.4}$$

There are four "corner" fixed points, (0,0), (1,0), (0,1), and (1,1). When $x = 0$, the trajectory is confined to the left edge of the state space, and converges to the equilibrium (0,0). When $x = 1$, it is on the right edge and converges to (1,1). When $n = 0$, the dynamics follow a replicator dynamic corresponding to the base game A_0, and when $n = 1$, the dynamic converges to (0,1) since this corresponds to replicator

dynamics of the PD game. However, because the set $(0,1)^2$ is an invariant set of these dynamics, we focus our attention on stability properties of trajectories starting in the interior of the state space.

4.2.3 Summary of possible dynamics in feedback-evolving games

The asymptotic behavior of the system (4.4) relies on the choice of the payoff parameters R_0, S_0, T_0, and P_0 of the game A_0. There are seven possible dynamical regimes, and they are summarized and named in Figure 4.2. The outcomes that are possible include a tragedy of the commons (TOC1 - TOC4), aversion of TOC (V1 and V2), and an "oscillating" TOC (OTOC). We present here analysis of one particular dynamical regime of note, where the system exhibits periodic orbits or asymptotically stable heteroclinic cycles. These dynamical outcomes are termed "oscillating tragedy of the commons" [119] because they are characterized by cycles between replete and deplete environmental states. The incentive structures required for these outcomes are stated in the following results.

Proposition 4.1. *Suppose $S_0 > P_0$ and $R_0 > T_0$. That is, mutual cooperation is the unique Nash equilibrium of the game A_0. Then the system* (4.4) *exhibits periodic orbits when*

$$\frac{P_1 - S_1}{T_1 - R_1} = \frac{S_0 - P_0}{R_0 - T_0} \tag{4.5}$$

In Figure 4.2, these payoffs correspond to the points on the line that separate the OTOC and V2 regions. The trajectories of the dynamics in this regime are shown in Figure 4.3.

Proposition 4.2. *Suppose $S_0 > P_0$ and $R_0 > T_0$. That is, mutual cooperation is the unique Nash equilibrium of the game A_0. Then the heteroclinic cycle Λ, defined by the orbits $(0,0) \to (1,0) \to (1,1) \to (0,1) \to (0,0)$, is asymptotically stable when*

$$\frac{P_1 - S_1}{T_1 - R_1} > \frac{S_0 - P_0}{R_0 - T_0} \tag{4.6}$$

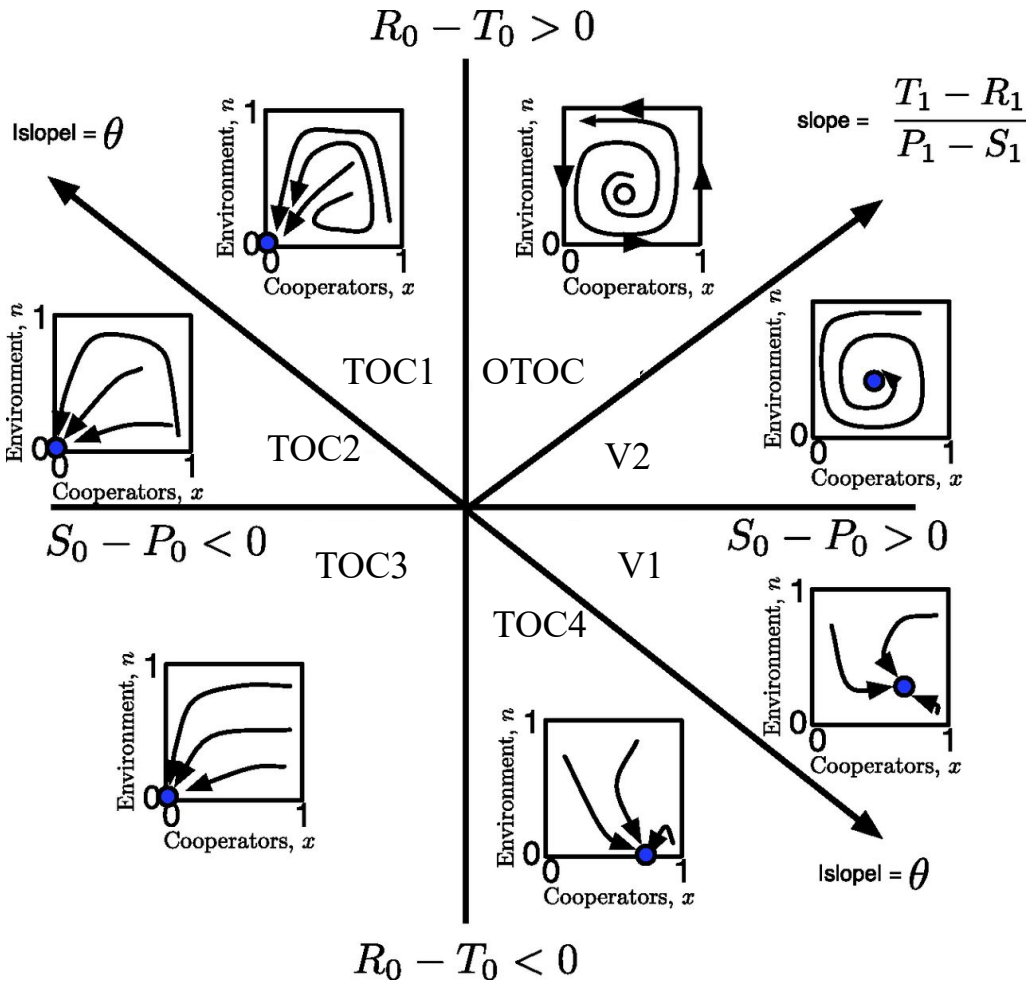


Figure 4.2: (Adapted from [119]) Summary of all possible dynamical outcomes given choice of payoffs in deplete state. The regions are determined by the relative payoffs $S_0 - P_0$ (x-axis) and $R_0 - T_0$ (y-axis). The phase portraits are illustrated in each region, where blue dots indicate stable fixed points of the dynamics. The seven regions include outcomes where a tragedy of the commons (TOC) occurs, and where a TOC is averted. We assign labels to each region, which includes four TOC outcomes, two averted outcomes (V1 and V2), and one oscillating TOC.

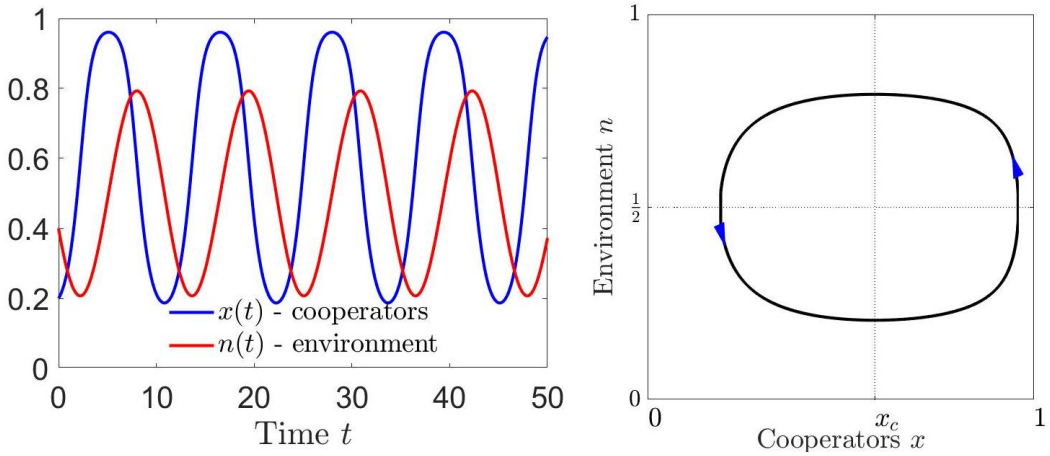


Figure 4.3: Solutions of $x(t)$, $n(t)$ in the regime specified by Proposition 4.1. (Left) Time series portrayal shows persistent oscillations. (Right) The $(x(t), y(t))$ trajectory traces out a closed orbit in the phase plane. Here, $[R_0, S_0, T_0, P_0] = [6, 4, 3, 3]$, $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, $\theta = 0.7$, and $[x_0, n_0] = [0.2, 0.4]$.

These payoffs correspond to the region OTOC in Figure 4.2. The trajectories of the dynamics in this regime are shown in Figure 4.4. A heteroclinic cycle is a set of equilibrium points connected by heteroclinic orbits, i.e. a path connecting two equilibrium points. In this system, the heteroclinic cyle is defined by the four corner equilibria points connected in the following order: (0,0), (1,0), (1,1), (1,0), and returning to (0,0). When such a cycle is asymptotically stable, state trajectories spend longer and longer times near the equilibrium points before going away to the next one in the cycle. We provide the proofs of these Propositions in Appendix 4.A. The stability properties of fixed points in the other dynamical regimes, along with their proofs, are also presented there.

4.3 A model of public opinion about the environment

Here, we introduce a model where the population responds to a public opinion of the environment instead of the true environmental state. Using this model, we investigate information control control policies aimed to influence the public opinion in order to maximally conserve the environmental state.

We build upon the feedback-evolving game model by introducing a public opinion

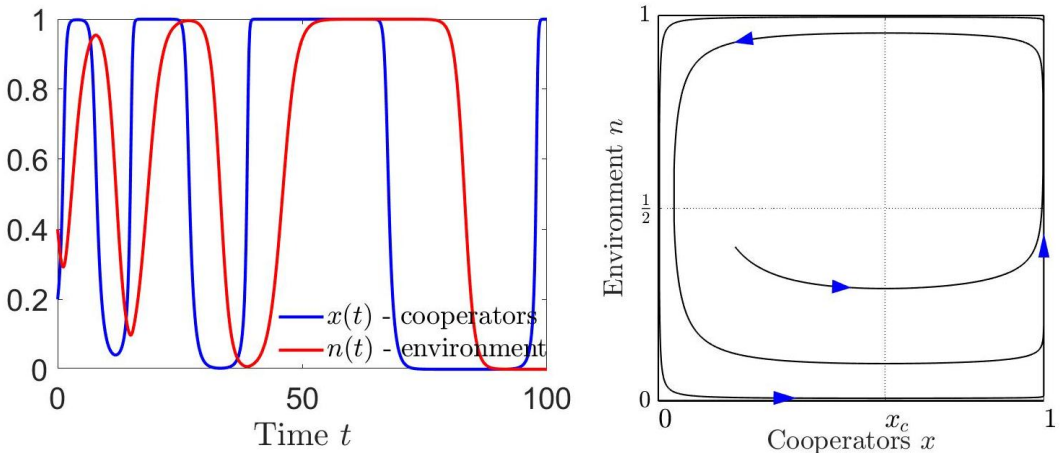


Figure 4.4: Solutions of $x(t)$, $n(t)$ in the regime specified by Proposition 4.2. (Left) Time series portrayal shows growing oscillations. (Right) The $(x(t), y(t))$ trajectory spirals outward towards the boundary of the state space, which defines a heteroclinic cycle. Here, $[R_0, S_0, T_0, P_0] = [10, 4, 3, 3]$, $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, $\theta = 0.7$, and $[x_0, n_0] = [0.2, 0.4]$.

variable $o(t)$. This variable is interpreted to be the average public opinion about the environmental state in the population, and it is a value between 0 and 1. For $o = 0$, all individuals in the population believe the environment is completely depleted, and when $o = 1$, they believe the environment is at a fully replete state. We introduce the following dynamics to model how opinions change in the population.

$$\dot{x} = x(1-x)g(x,o)$$

$$\dot{n} = n(1-n)(-1+(1+\theta)x)$$

$$\dot{o} = -\gamma(o-n)$$

$$[x(0), n(0), o(0)] = [x_0, n_0, o_0] \in (0,1)^3$$

where $\gamma > 0$. The form of the $\dot{o}$ equation induces $o(t)$ to track the environmental state $n(t)$. There is a lag between actual changes in the environment and the public becoming informed about the changes. The learning parameter $\gamma > 0$ determines how slow this lag is. For γ low, $o(t)$ will not adapt quickly enough to the fluctuating $n(t)$. As γ increases, $o(t)$ more successfully tracks $n(t)$. The $\dot{x}$ equation above is modified from the previous section by replacing the relative fitness $g(x,n)$ with $g(x,o)$. Here,

individual incentives are now determined by the current public opinion and not the true environmental state n. We denote this system with the mapping F^o, which is written

$$\dot{\boldsymbol{y}} \equiv F^o(\boldsymbol{y}) = \begin{bmatrix} x(1-x)g(x,o) \\ n(1-n)(-1+(1+\theta)x) \\ -\gamma(o-n) \end{bmatrix}$$

where $\boldsymbol{y} = [x, n, o]^\top$ is the state vector. The previous feedback-evolving game dynamics F (4.4) can be interpreted here as a population that has perfect knowledge of the environmental state. We illustrate the effect delayed public opinion has on the original dynamics in Figure 4.5 in three distinct regimes. A notable effect occurs in the V2 regime, where the original dynamics admits a stable interior fixed point at $(1/(1+\theta), n_c)$, where

$$n_c \equiv \frac{(R_0 - T_0) + \theta(S_0 - P_0)}{(R_0 - T_0) + \theta(S_0 - P_0) + (T_1 - R_1) + \theta(P_1 - S_1)}.$$

With public opinion, the $(x(t), n(t))$ trajectories of the system F^o are pushed towards the boundary of the state space $(x, n) \in [0, 1]^2$. This is due to the delay in opinion, and the intuition is as follows. When $n(t)$ starts to increase towards a peak, $o(t)$ lags behind and stays below $n(t)$. This causes more of the population to become cooperators, since they are responding to the low public opinion and not the true environmental state. As a result, $n(t)$ is restored more than it would have been if the population had perfect information. As $o(t)$ tracks increases in $n(t)$, the population starts to defect, degrading the environment. Opinion starts to decline as well, causing a resurgence in cooperation. This process continues to repeat, causing larger and larger oscillations. The delay of opinion causes an imbalance in population response because individuals are responding to exaggerated opinions of the environment.

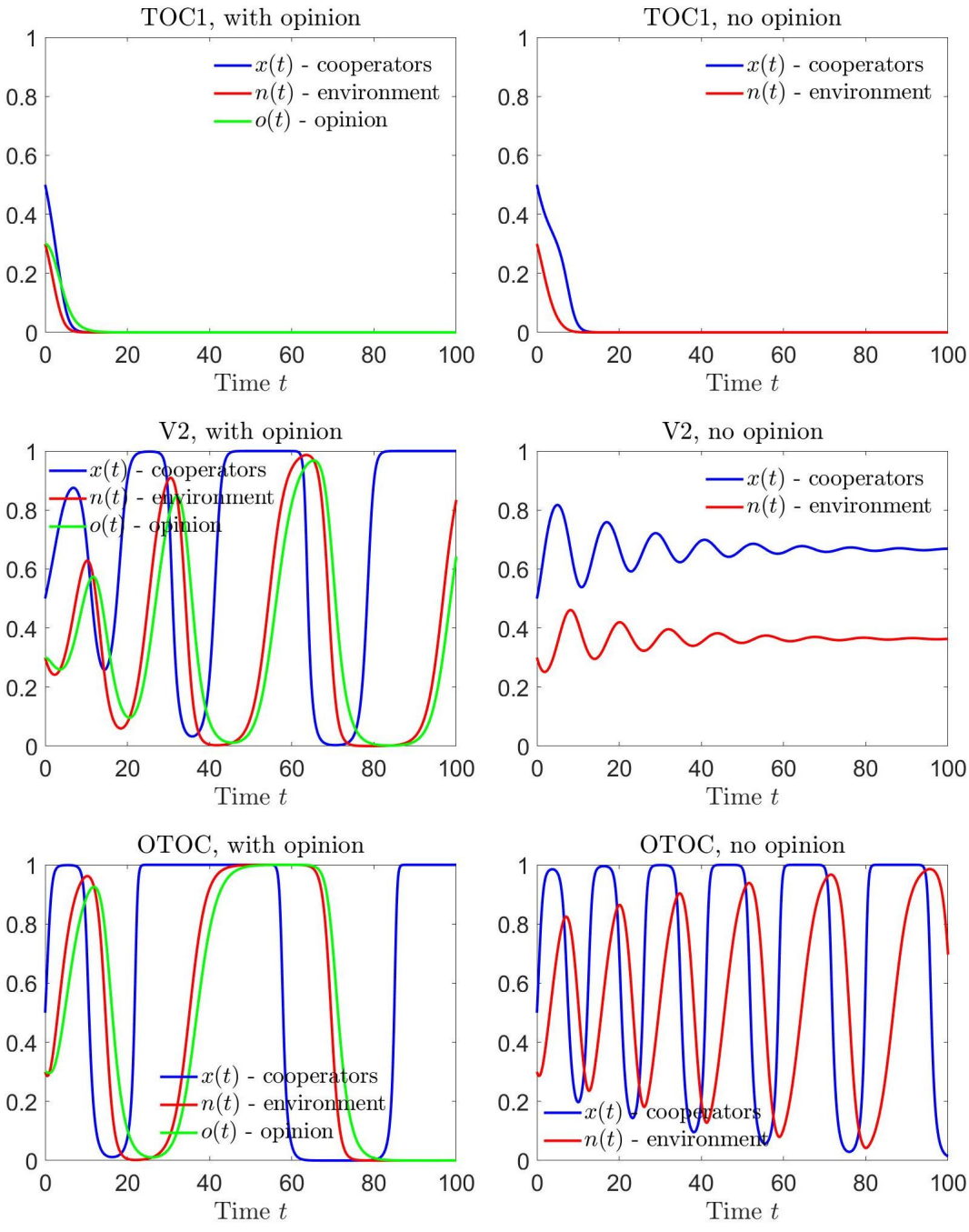


Figure 4.5: Comparison of public opinion-induced dynamics (Left column) and the original feedback-evolving game (Right column). (Top row) $[R_0, S_0, T_0, P_0] = [5, 2, 3, 3]$ (regime TOC1). Delay of opinion does not help to restore the commons. (Middle row) $[R_0, S_0, T_0, P_0] = [4.5, 4, 3, 3]$ (regime V2). Delayed opinion destabilizes the interior fixed point. (Bottom row) $[R_0, S_0, T_0, P_0] = [7, 4, 3, 3]$ (regime OTOC). Public opinion facilitates convergence to heteroclinic cycle in the (x, n) trajectories. In all simulations, $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, $\gamma = .5$, $\theta = .5$, $x_0 = .5$, $n_0 = .3$, $o_0 = .3$.

4.4 Opinion influence policies for conservation

Using the population opinion model of the previous section, in this section we consider an external entity, e.g. media platforms, politicians, activists, and central figures, that has an influence over public opinion about the environment. We numerically characterize control policies intended to influence public opinion with the goal of conserving the environmental state by utilizing an optimal control algorithm. We consider two types of information control policies by formulating two optimal control problems. First, we study strategic policies that perturb public opinions externally. These types of strategies inject information to sway public opinion. For example, propaganda, activist campaigns, and media broadcasts can achieve such perturbations. We then consider policies that facilitate public learning of the true environmental state. These strategies raise awareness about the environment and serve to reduce uncertainty about the environmental state. For example, the EPA conducts environmental education programs[2].

Before proceeding, we outline a Hamiltonian-based descent algorithm (Algorithm 3.3 of [39]) that we utilize in both optimal control problems. The algorithm serves to numerically compute a control function $u(t)$ that locally maximizes the objective function J, which we will define for both problems to be

$$\max_u J(u) \equiv \int_0^{T_f} L(n,u)dt = \frac{1}{2}\int_0^{T_f} C_1 n^2 - C_2 u^2 dt$$
$$\text{subject to } \dot{\boldsymbol{y}} = F^o(\boldsymbol{y}, u)$$

where $L(n,u) \equiv \frac{1}{2}(C_1 n^2 - C_2 u^2)$ is the *running cost*, $T_f > 0$ is the finite time horizon, and $C_1, C_2 > 0$ are weighting constants. The $C_1 n^2$ term rewards accumulation of the environmental state. The $-C_2 u^2$ term penalizes control effort. The controlled state dynamics $F^o(\boldsymbol{y}, u)$ will be specified for the two problems we consider. Both problems

[2]https://www.epa.gov/education/what-environmental-education

will be in Bolza form with no terminal cost. The Hamiltonian is

$$H(\boldsymbol{y}, \boldsymbol{\lambda}, u) = \boldsymbol{\lambda}^\top F^o(\boldsymbol{y}, u) + L(n, u)$$

where $F^o(\boldsymbol{y}, u)$, the controlled state dynamics, will be different for the two problems we consider. Here, $\boldsymbol{\lambda} = [\lambda_x, \lambda_n, \lambda_o]^\top$ are the costate variables. The costates are determined by the dynamical equations

$$\begin{aligned} \dot{\lambda}_x &= -\frac{\partial H}{\partial x}(\boldsymbol{y}, \boldsymbol{\lambda}, u) \\ \dot{\lambda}_n &= -\frac{\partial H}{\partial n}(\boldsymbol{y}, \boldsymbol{\lambda}, u) \\ \dot{\lambda}_o &= -\frac{\partial H}{\partial o}(\boldsymbol{y}, \boldsymbol{\lambda}, u) \\ &\text{with final condition } \boldsymbol{\lambda}(T_f) = [0, 0, 0]^\top \end{aligned} \tag{4.7}$$

The final conditions are all zero because of the absence of a terminal cost. The iterative algorithm of [39] leverages Pontryagin's Maximum Principle and the Hamiltonian to find ascent directions for the control function. The algorithm operates as follows.

Algorithm 4.1. Set the number of iterations `iters`, a positive integer. Fix constants $\alpha \in (0, 1)$ and $\beta \in (0, 1)$. These are used for the Armijo step search [7].

1. Start with an initial guess for the control function, u_0, and set $k = 0$.

2. At iteration k, compute the state trajectory $\boldsymbol{y}_k$ by solving the system dynamics $F^o(\boldsymbol{y}, u_k)$ using a forward integration method. Then, using these trajectories and the current control iterate u_k, compute the costate trajectories $\boldsymbol{\lambda_k}$ (4.7) by using a backward integration method.

3. Compute the pointwise (for all $t \in [0, T_f]$) maximizer $u_k^* = \arg\max_u H(\boldsymbol{y}_k, \boldsymbol{\lambda}_k, u)$ of the Hamiltonian.

4. (Armijo step search) Compute the smallest integer $\ell = 0, 1, \ldots$, such that

$$J(u_k) - J(u_k + \beta^\ell (u_k^* - u_k)) \leq \alpha \beta^\ell \Theta(u_k)$$

where $\Theta(u)$ is the *optimality function*, which is always non-positive.

$$\Theta(u) = \int_0^{T_f} (H(\boldsymbol{y}, \boldsymbol{\lambda}, u) - H(\boldsymbol{y}, \boldsymbol{\lambda}, u^*))dt \leq 0$$

where the trajectories $\boldsymbol{y}$ and $\boldsymbol{\lambda}$ correspond to the control u.

5. Update

$$u_{k+1} = u_k + \beta^\ell (u^* - u_k)$$

and $k \leftarrow k + 1$. If $k >$ `iters`, exit algorithm. If not, return to step 2.

The value of $|\Theta(u_k)|$ measures how far from optimality the current iterate is, in the sense of Pontryagin's Maximum Principle. When it is zero, the conditions of Pontryagin's Maximum Principle are met.

4.4.1 Influence through propaganda

We consider strategic policies that perturb public opinions externally by introducing new, artificial information into the population. To this end, we formulate the following finite time horizon (T_f) optimal control problem, with $u(t) \in \mathbb{R}$ for all $t \in [0, T_f]$ as a scalar-valued control function.

$$\begin{aligned} &\max_u J = \frac{1}{2} \int_0^{T_f} C_1 n^2 - C_2 u^2 dt \\ &\text{subject to} \\ &\dot{x} = x(1-x)g(x,o) \\ &\dot{n} = n(1-n)(-1 + (1+\theta)x) \\ &\dot{o} = -\gamma(o-n) + o(1-o)u \end{aligned} \tag{4.8}$$

The additive control term $o(1-o)u$ serves two purposes. First, it keeps the dynamics well-posed, i.e. a solution $o(t)$ that starts in $[0, 1]$ will stay in $[0, 1]$. Second, it models the difficulty to influence extreme opinions. The additive term decreases to zero as o approaches the extremes 0 and 1, and hence more external influence is required to move public opinion away from the extremes. The system dynamics of the above are

denoted here as $\dot{\boldsymbol{y}} = F^o(\boldsymbol{y}, u)$. An information control policy must balance the costs of spreading misinformation, e.g. public broadcasts, with its benefits.

To implement Algorithm 4.1, we first need to compute the co-state equations

$$\begin{aligned}
\dot{\lambda}_x &= -\lambda_x \left((1-2x)g(x,o) + x(1-x)\frac{\partial g}{\partial x}(o) \right) - \lambda_n n(1-n)(1+\theta) \\
\dot{\lambda}_n &= -\lambda_n(1-2n)(-1+(1+\theta)x) - \lambda_o\gamma - C_1 n \\
\dot{\lambda}_o &= -\lambda_x x(1-x)\frac{\partial g}{\partial o}(x) - \lambda_o(-\gamma + u(1-2o)) \\
&\quad \text{with final condition } \boldsymbol{\lambda}(T_f) = [0,0,0]^\top
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial g}{\partial x}(o) &= (1-o)((R_0-T_0)+(P_0-S_0)) + o((R_1-T_1)+(P_1-S_1)) \\
\frac{\partial g}{\partial o}(x) &= (1-x)((P_0-S_0)-(P_1-S_1)) + x((R_1-T_1)-(R_0-T_0)).
\end{aligned}$$

The Hamiltonian can be written

$$\begin{aligned}
H(\boldsymbol{y}, \boldsymbol{\lambda}, u) &= \boldsymbol{\lambda}^\top \dot{\boldsymbol{y}} + L(n,u) \\
&= \lambda_x x(1-x)g(x,o) + \lambda_n n(1-n)(-1+(1+\theta)x) + \lambda_o(-\gamma(o-n) + o(1-o)u) + \cdots \\
&\quad + \frac{1}{2}(C_1 n^2 - C_2 u^2).
\end{aligned}$$

The above expression is concave in u. Hence, the point-wise (in time) maximizer of the Hamiltonian is calculated by finding where the partial derivative of H with respect to u is zero. We obtain

$$u^*(t) = \frac{1}{C_2}\lambda_o(t)o(t)(1-o(t))$$

We applied Algorithm 4.1 to the control problem (4.8). We initialized all simulations with the initial control guess $u(t) = 0$ for all $t \in [0, T_f]$, fixed payoffs $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, and initial conditions $[x_0, n_0, o_0]^\top = [.5, .3, .3]$. We utilized RK4-based integration methods (ode45 for forward integration and RK4 for backwards), with uniform spacing .001 for the time variable.

We characterize the efficacy of information injection policies through a suite of simulations applying the control algorithm to the different dynamical regimes. In all of the simulations we conduct, we fix the priority weight $C_1 = 1$, and study modifications to the penalty weight C_2. First, a notable observation was that the environmental state could be rescued in the TOC1 regime for a limited time if penalty on effort was extremely low ($C_2 = .001$). When the cost weights are more balanced, e.g. $C_2 = 1$, we did not observe resurgence of the commons in any of the TOC regimes after implementation of Algorithm 4.1. In the other regimes, the control applies its effort into pushing $o(t)$ lower. In the cases we have studied, the algorithm produces controls $u(t)$ that are negative throughout the time horizon.

In regime V2 (Figure 4.6), when control effort is cheap ($C_2 = 0.1$), the computed control concentrates effort in waves, occurring near the peaks of the oscillations. The control starts very negative, leading to a decrease in $o(t)$ in the first few time steps. This produces a trough in the $o(t)$ oscillation that is slightly lower than $n(t)$, resulting in a resurgence of cooperators. The control then relaxes its effort as the states $o(t), n(t)$ begin to increase, approaching the next peak (time 10). At this peak, information is applied to drive $o(t)$ higher, resulting in a resurgence of defectors. This process is repeated again for the next cycle, causing larger oscillations in the state variables.

In the OTOC regime, control effort is applied in a similar manner, but with more selective pushes. In Figure 4.7 (Top right), there are two concentrated efforts to push $o(t)$ lower at $t \approx 5$ and $t \approx 30$. Both pushes are followed by relaxations in effort, returning to $u \approx 0$. This control intervention facilitates an oscillating tragedy of the commons. However, the magnitudes of these pushes are small due to the higher penalty $C_2 = 1$. When the penalty is relaxed to $C_2 = 0.001$ (Figure 4.8), all control effort is concentrated at the beginning, and immediately relaxes to near zero for the rest of the time horizon. This causes $o(t)$ to shoot down, leading to a sudden surge

in cooperators and consequently restoration of the environment. The relaxation of the control causes $o(t)$ to eventually track the high value of $n(t)$, leading to a sudden resurgence in defectors at $t \approx 40$. Beyond the time horizon $T_f = 50$ and in the absence of additional control, the defector dominance will eventually lead to a collapse of the environmental state.

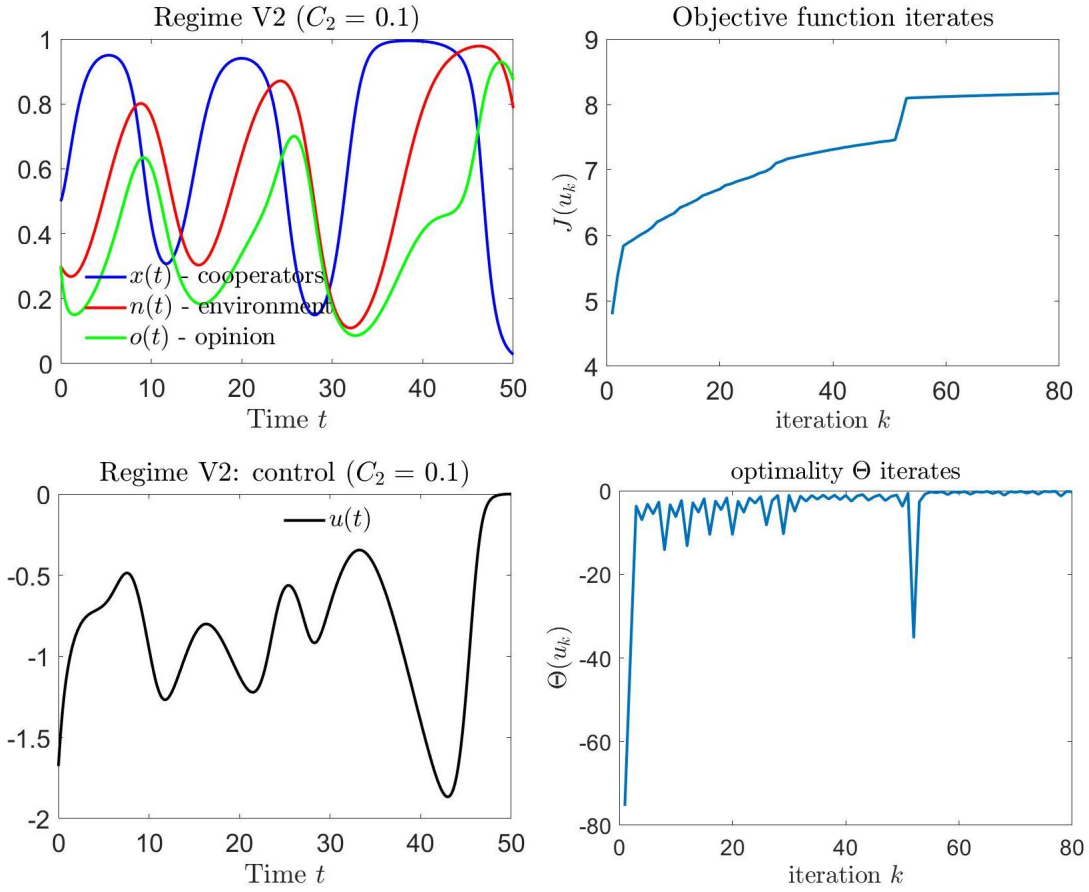


Figure 4.6: Results for information injection problem in the V2 regime (stable interior fixed point) with effort penalty $C_2 = 0.1$. The initial control guess was $u(t) = 0$, $[R_0, S_0, T_0, P_0] = [4.5, 4, 3, 3]$, $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, $\gamma = 1$, $\theta = .5$, $x_0 = .5$, $n_0 = .3$, $o_0 = .3$, $T_f = 50$. We ran the algorithm for 80 iterations. (Top left) The controlled state dynamics produce an oscillating tragedy of the commons. (Top right) The objective function J approaches saturation. (Bottom left) The control function at iteration 80. (Bottom right) Optimality function converges to near zero, indicating a local maximum.

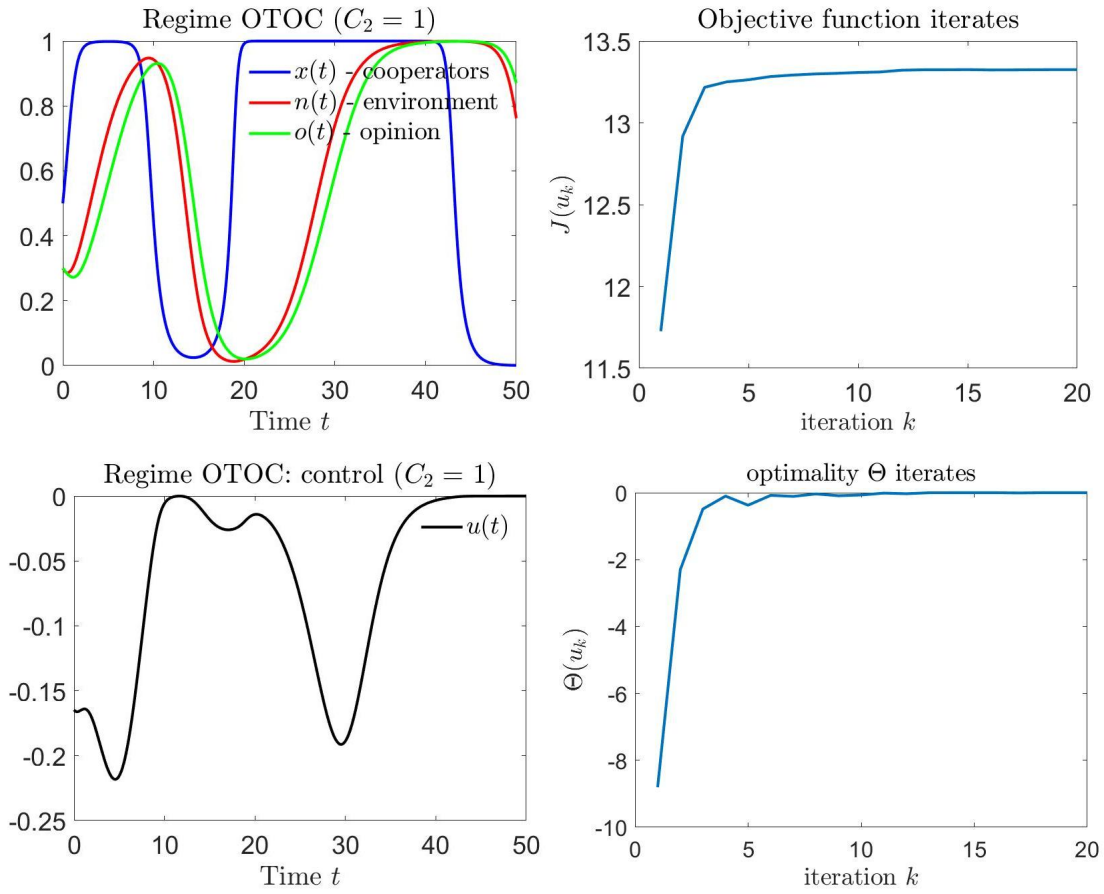


Figure 4.7: Results for information injection problem in the OTOC regime (stable heteroclinic cycle) with effort penalty $C_2 = 1$. The initial control guess was $u(t) = 0$, $[R_0, S_0, T_0, P_0] = [7, 4, 3, 3]$, $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, $\gamma = 1$, $\theta = .5$, $x_0 = .5$, $n_0 = .3$, $o_0 = .3$, $T_f = 50$. We ran the algorithm for 20 iterations. (Top left) With additive control, state dynamics produce an oscillating tragedy of the commons. (Top right) The objective function J approaches saturation. (Bottom left) The control function at iteration 20. (Bottom right) Optimality function converges to near zero, indicating a local maximum.

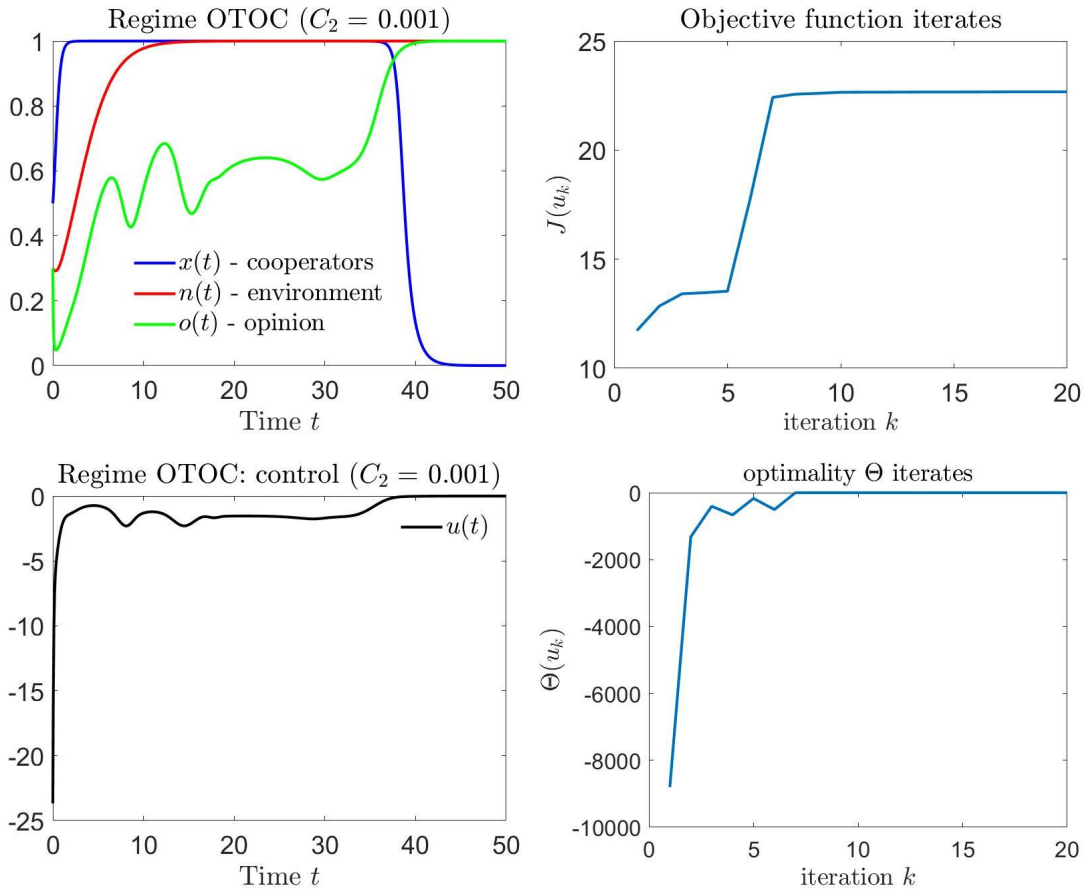


Figure 4.8: Results for information injection problem in the OTOC regime (stable heteroclinic cycle) with effort penalty $C_2 = .001$. The initial control guess was $u(t) = 0$, $[R_0, S_0, T_0, P_0] = [7, 4, 3, 3]$, $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, $\gamma = 1$, $\theta = .5$, $x_0 = .5$, $n_0 = .3$, $o_0 = .3$, $T_f = 50$. We ran the algorithm for 20 iterations. (Top left) With additive control, state dynamics produce an oscillating tragedy of the commons. (Top right) The objective function J approaches saturation. (Bottom left) The control function at iteration 20. (Bottom right) The value of the optimality function at iteration 80 is -0.008347, indicating a local maximum.

4.4.2 Influence through raising awareness

We now consider strategic information policies that guide public opinion towards the true environmental state. Environmental awareness and educational campaigns are examples of interventions that serve this purpose. Using Algorithm 4.1, we wish to numerically characterize how such policies are should be applied to conserve the environmental state. We formulate the following finite time-horizon optimal control problem, with $u(t) \in [0, \infty)$ for all $t \in [0, T_f]$ as a scalar-valued control function.

$$\max_u J = \frac{1}{2} \int_0^{T_f} C_1 n^2 - C_2 u^2 dt$$

subject to

$$\begin{aligned} & u(t) \in [0, \infty) \ \forall t \in [0, T_f] \\ & \dot{x} = x(1-x)g(x, o) \\ & \dot{n} = n(1-n)(-1 + (1+\theta)x) \\ & \dot{o} = -(\gamma + u)(o - n) \end{aligned} \tag{4.9}$$

In this formulation, we have constrained the control to be non-negative for all times. This serves two purposes. First, it ensures the dynamics will be well-posed, i.e. all states stay within the interval $(0, 1)$. Second, it restricts the influencing entity from mis-educating the public since the learning parameter $\gamma + u(t) \geq \gamma$ for all t.

To implement Algorithm 4.1, we need to calculate the costate dynamics, which are

$$\begin{aligned} \dot{\lambda}_x &= -\lambda_x \left((1-2x)g(x, o) + x(1-x)\frac{\partial g}{\partial x}(o) \right) - \lambda_n n(1-n)(1+\theta) \\ \dot{\lambda}_n &= -\lambda_n (1-2n)(-1 + (1+\theta)x) - \lambda_o(\gamma + u) - C_1 n \\ \dot{\lambda}_o &= -\lambda_x x(1-x)\frac{\partial g}{\partial o}(x) + \lambda_o(\gamma + u) \end{aligned}$$

with final condition $\boldsymbol{\lambda}(T_f) = [0, 0, 0]^\top$. The Hamiltonian is

$$\begin{aligned} H(\boldsymbol{y}, \boldsymbol{\lambda}, u) = \lambda_x x(1-x)g(x, o) + \lambda_n n(1-n)(-1 + (1+\theta)x) - \lambda_o(\gamma + u)(o - n) + \cdots \\ + \frac{1}{2}(C_1 n^2 - C_2 u^2). \end{aligned}$$

The above expression is concave in u, where the relevant term is $-u(\lambda_o(o-n) + (1/2)C_2u)$. If the maximizer u^* of this term is positive, then the maximizer takes that value. If it is negative or zero, our constraint restricts u^* to zero. Hence, the point-wise (in time) maximizer of the Hamiltonian is written as

$$u^*(t) = \arg\max_{u\in[0,\infty)} H(\boldsymbol{y}, \boldsymbol{\lambda}, u) = \begin{cases} 0 & \text{if} -(1/C_2)\lambda_o(o-n) < 0 \\ -(1/C_2)\lambda_o(o-n) & \text{if } -(1/C_2)\lambda_o(o-n) \geq 0 \end{cases}$$

We applied Algorithm 4.1 to the control problem (4.9). In these simulations, we lower $\gamma = 0.5$ to increase delay between public opinion $o(t)$ and enviornmental state $n(t)$. We find that even with low penalty on control effort $C_2 = .001$ and priority on conservation $C_1 = 1$, the computed controls were unable to induce a resurgence of the commons in all TOC regimes. In the other regimes, the control applies public opinion guidance effort near critical points in the time horizon. To illustrate these effects, we present the outcomes in the dynamical regimes V2 and OTOC. We initialized all simulations with the initial control guess $u(t) = 0$ for all $t \in [0, T_f]$, fixed payoffs $[R_1, S_1, T_1, P_1] = [3, 1, 6, 2]$, initial conditions $[x_0, n_0, o_0]^\top = [.5, .3, .3]$, selected the learning paramter $\gamma = 0.5$, and fixed conservation priority $C_1 = 1$.

We characterize the efficacy of awareness policies in the problem (4.9) through a suite of simulations applying Algorithm 4.1 to the different dynamical regimes. In all of the simulations we conduct, we fix the priority weight $C_1 = 1$, and study modifications to the penalty weight C_2. We find that even with low penalty on control effort $C_2 = .001$, the computed awareness policies were unable to induce a resurgence of the commons in all TOC regimes.

In regime V2 (Figure 4.9), we set an inexpensive control effort ($C_2 = .001$). The computed control (Bottom left) raises public awareness of the true environmental state only at a few critical moments in the time horizon. These moments occur near the first few extrema of the $n(t)$ oscillations. Because opinion lags the true state, $o(t)$

underestimates $n(t)$ when they are increasing and overestimates $n(t)$ when decreasing. The sudden injection of awareness allows $o(t)$ to quickly track the more extreme value of $n(t)$ near the critical points, inducing higher and higher oscillations of $o(t)$. This effect is most notably seen in the first peak of $o(t)$ (Top left).

Figure 4.10 shows an application of the algorithm in the OTOC regime. Here we also set $C_2 = .001$. A similar principle holds for the computed controller here as well. An initial spike in influence causes opinion to quickly shoot up to meet the higher $n(t)$. The controller immediately relaxes, causing $o(t)$ to overestimate $n(t)$ as it decreases. A second spike at $t \approx 7.5$ causes opinion to shoot down to meet the lower $n(t)$, now at a trough. No more control is applied for the rest of the horizon, but the two perturbations cause $o(t)$ to approach more extreme values in its oscillations (Left panels).

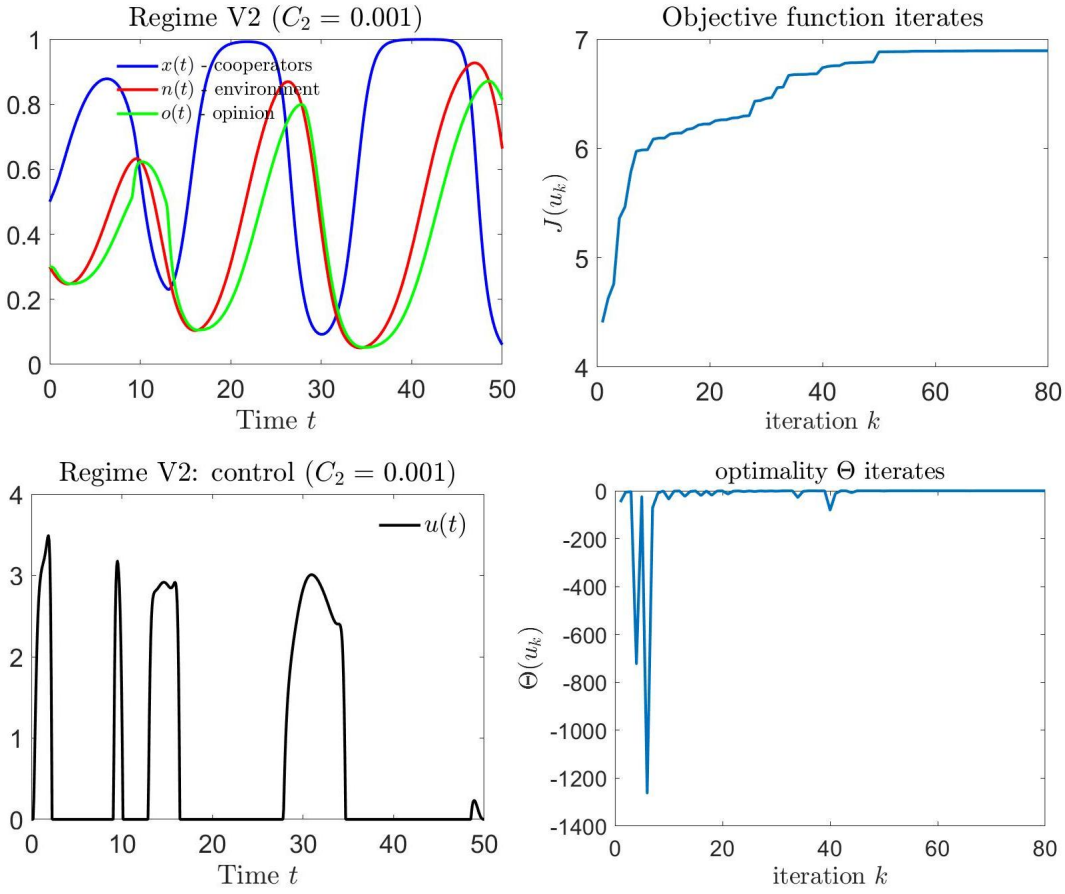


Figure 4.9: Simulation results for the awareness control problem in the V2 regime (stable interior fixed point) with very low penalization $C_2 = .001$. Here, $[R_0, S_0, T_0, P_0] = [4.5, 4, 3, 3]$, $\gamma = 0.5$, $\theta = .5$, $x_0 = .5$, $n_0 = .3$, $o_0 = .3$, $T_f = 50$. (Top left) The controlled states resemble an oscillating tragedy of the commons. (Top right) The objective function J approaches saturation after 80 iterations. (Bottom left) The control function at iteration 80. (Bottom right) The value of the optimality function at iteration 80 is $\approx -4 \times 10^{-5}$, indicating a local maximum.

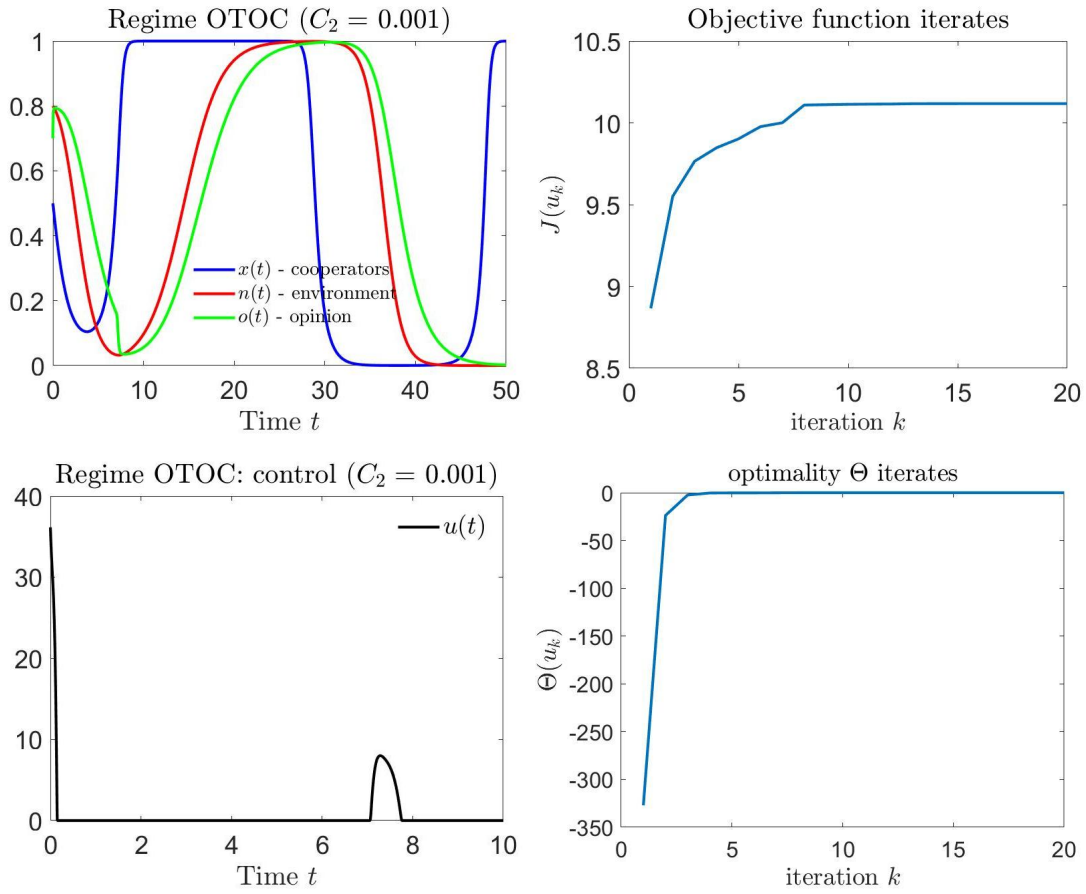


Figure 4.10: Simulation results for the awareness control problem in the OTOC regime (stable heteroclinic cycle) with very low penalization $C_2 = .001$. Here, $[R_0, S_0, T_0, P_0] = [7, 4, 3, 3]$, $\gamma = 0.5$, $\theta = .5$, $x_0 = .5$, $n_0 = .8$, $o_0 = .7$, $T_f = 50$. (Top left) The controlled states resemble an oscillating tragedy of the commons. (Top right) The objective function J approaches saturation. (Bottom left) The control function at iteration 80. Here it is displayed on $[0, 10]$ to magnify the shape of the first spike. The control $u(t)$ is also zero for the rest of the horizon. (Bottom right) The value of Θ at iteration 80 here is $\approx -4.29 \times 10^{-5}$, indicating a local maximum.

4.5 Incentive policies for conservation

In this section, we consider a different type of strategic policy - one that influences individuals' incentives to cooperate together with the goal of conserving public resources. For example, government entities may incentivize fishers to refrain from fishing activity through buyouts. The control variable is applied in the payoff matrix as follows.

$$A_n(u(t)) = n \begin{bmatrix} R_1 & S_1 \\ T_1 & P_1 \end{bmatrix} + (1-n) \begin{bmatrix} R_0 + u(t) & S_0 \\ T_0 & P_0 \end{bmatrix}$$

We will constrain $u(t) \in [-u_m, u_m] \equiv U$ for all t, where $u_m > 0$ is a positive constant. Hence, we formulate this as the following optimal control problem.

$$\begin{aligned} &\max_{u \in U} J = \int_0^{T_f} n^2(t)dt \\ &\text{subject to } \begin{cases} \dot{x} = x(1-x)g(x,n) + x^2(1-x)(1-n)u \\ \dot{n} = n(1-n)(-1+(1+\theta)x) \\ x(0) = x_0 \in [0,1],\ n(0) = n_0 \in [0,1] \end{cases} \end{aligned} \tag{4.10}$$

Note that we are here using the original feedback-evolving game (without opinions). The affine term $x^2(1-x)(1-n)$ appears after re-deriving the replicator equation with the payoff matrix $A_n(u)$. Again we are able to apply Algorithm 4.1 to (4.10) by first computing the costates

$$\dot{\lambda_x} = -\lambda_x \left((1-2x)g(x,n) + x(1-x)\frac{\partial g(n)}{\partial x} + x(2-3x)(1-n)u \right) - \lambda_n n(1-n)(1+\kappa)$$
$$\dot{\lambda_n} = -\lambda_x \left(x(1-x)\frac{\partial g(x)}{\partial n} - x^2(1-x)u \right) - \lambda_n(1-2n)(-1+(1+\kappa)x) - 2n$$

and the Hamiltonian

$$H(\boldsymbol{y}, \boldsymbol{\lambda}, u) = \lambda_x x(1-x)(g(x,n) + x(1-n)u) + \lambda_n n(1-n)(-1+(1+\kappa)) + n^2 \tag{4.11}$$

Because the affine term $x^2(1-x)(1-n)$ is always positive, the pointwise maximizer of the Hamiltonian is

$$u^*(t) = u_m \text{sgn}(\lambda_x(t)) \tag{4.12}$$

The maximizer $u^*(t)$ will only take two values, the minimum and maximum points in the constraint set $U = [-u_m, u_m]$. As long as $\lambda_x(t) = 0$ does not occur on an open interval in the time horizon $[0, T_f]$, $u^*(t)$ is classified as a "bang-bang" controller. We will shortly prove that this is indeed true, using the Lie bracket (Ch. 4.4 of [55]).

We applied Algorithm 4.1 with the parameters $[R_0, S_0, T_0, P_0] = [4.5, 4, 3, 3]$, $[R, S, T, P] = [3, 1, 6, 2]$ (regime V2), $u_m = 1$, $T_f = 100$, $\theta = 0.7$, $[x_0, n_0] = [0.7, 0.3]^\top$. Shown in Figure 4.11 is the result of 40 iterations of the algorithm with initial control guess $u_0(t) = 0$. The computed control converges to a local maximum, indicated numerically by convergence of the optimality function Θ to zero. This controller appears to switch values near the times where $n(t)$ peaks. At these peaks, $x(t) = x_c = 1/(1+\theta)$. In doing so, the controlled dynamics exhibit growing oscillations, in a regime (V2) where uncontrolled dynamics settle at an interior fixed point.

We run a second simulation (Figure 4.12 with the same parameter set, but select the initial guess $u_0 = u_m \text{sgn}(x - x_c)$ because we suspect such a controller is effective at driving growing oscillations in the system. Again, the algorithm produces a local maximum, as indicated numerically by convergence of the optimality function Θ to zero. The resulting controller after 40 iterations has deviated only slightly away from u_0, and the improvement in performance by the end is only marginally improved (29.84 to 29.97). This controller outperforms the computed control from the first simulation, where the initial guess was the zero function.

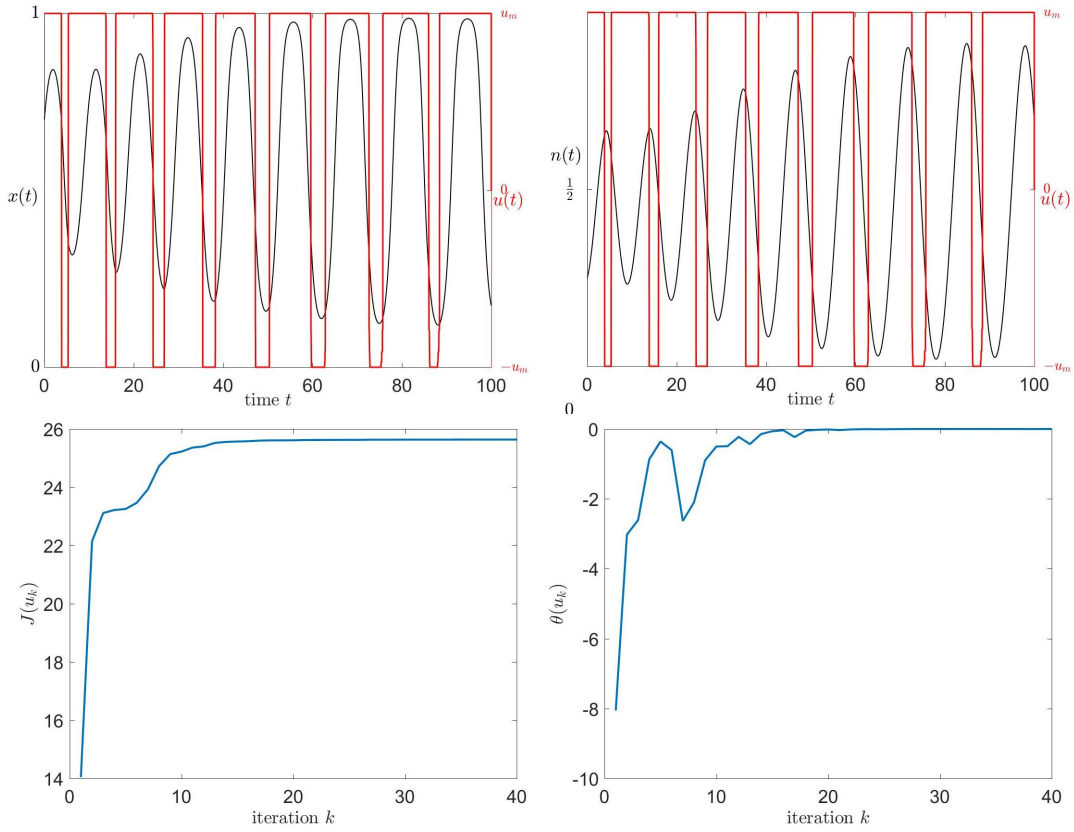


Figure 4.11: Result of Algorithm 4.1 applied to the incentive control problem. We set the initial guess $u_0(t) = 0$. At minimal and maximal control ($u = \pm 1$), the dynamic still remains in regime V2. (Top Left) The cooperator dynamics $x(t)$ overlayed with the computed control $u_{40}(t)$ at iteration 40. (Top Right) Environment dynamics $n(t)$. The control u_{40} switches near the critical times where $\dot{n}(t) = 0$. (Bottom Left) Objective scores $J(u_k) = \int_0^{T_f} n^2(t)dt$ versus iteration number k. The score approaches a maximal value (≈ 26). (Bottom Right) The optimality function $\theta(u_k)$ vs iteration number k.

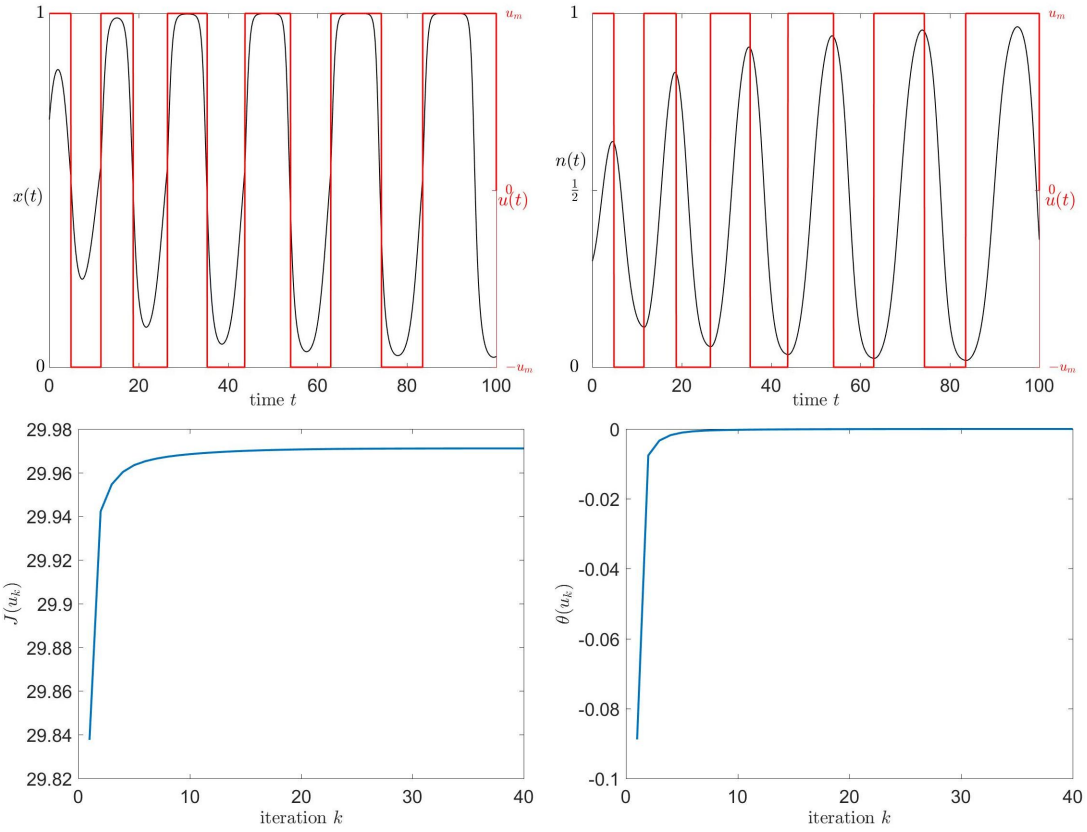


Figure 4.12: Same set-up as Figure 1, but with initial control guess $u_0 = u_m \text{sgn}(x - x_c)$. (Top Right) Environment dynamics $n(t)$. The control u_{40} switches near the critical points where $\dot{n}(t) = 0$. (Bottom Left) Objective scores $J(u_k) = \int_0^{T_f} n^2(t)dt$ versus iteration number k. The score approaches a maximal value (≈ 30), which is slightly higher than Figure 1 setup. Hence, we observe this local max performs better. (Bottom Right) The optimality function $\theta(u_k)$ vs iteration number k.

4.5.1 Lie bracket analysis - non-singular bang-bang incentive control

Here, we analyze the Lie bracket of the incentive-based controlled dynamics (4.10) to show that an optimal control u^* (4.12) that satisfies the conditions of Pontryagin's Maximum Principle is indeed a bang-bang controller. This amounts to showing that there are no singular arcs - i.e. there does not exist an open interval in which the switching function is zero. See Chapter 4.4.3 of [55] for the relevant theory. Our controlled system dynamics (4.10) are control-affine, and can be written in the form $\dot{\boldsymbol{y}} = F(\boldsymbol{y}) + G(\boldsymbol{y})u$ where $\boldsymbol{y} = [x, n]^\top$. The affine term is

$$G(x, n) = [x^2(1-x)(1-n), 0]^\top$$

The switching function in our control problem is

$$\varphi(t) \equiv \langle \boldsymbol{\lambda}, G(x, n) \rangle = \lambda_x x^2 (1-x)(1-n)$$

That is, the control that maximizes the Hamiltonian (4.11) is

$$u^*(t) = \begin{cases} u_m, & \text{if } \varphi(t) > 0 \\ \times & \text{if } \varphi(t) = 0 \\ -u_m, & \text{if } \varphi(t) < 0 \end{cases}$$

where $\times$ indicates u^* can take an arbitrary value. The time derivative of the switching function is $\dot{\varphi} = \langle \boldsymbol{\lambda}, (DG)F - (DF)G \rangle = \langle \boldsymbol{\lambda}, [F, G] \rangle$, where $[F, G] = (DG)F - (DF)G$ denotes the Lie bracket of the vector fields F and G. We calculate the Jacobians

$$DG = \begin{bmatrix} x(2-3x)(1-n) & -x^2(1-x) \\ 0 & 0 \end{bmatrix}$$

$$DF = \begin{bmatrix} (1-2x)g(x,n) + x(1-x)\frac{\partial g}{\partial x} & x(1-x)\frac{\partial g}{\partial n} \\ n(1-n)(1+\kappa) & (1-2n)(-1+(1+\kappa)x) \end{bmatrix}$$

We calculate the Lie bracket to be

$$[F,G] = x^2(1-x)(1-n)\begin{bmatrix} g(x,n)(2-3x) - n(-1+(1+\kappa)x) - \left((1-2x)g(x,n) + x(1-x)\frac{\partial g}{\partial x}\right) \\ -n(1-n)(1+\kappa) \end{bmatrix}$$

We can rule out singularity of u^* ($\varphi(t) = 0$ on an open interval) if G and $[F,G]$ are linearly independent. Since $(x,n) \in (0,1)^2$ (invariance of the interior of state space) for all t, we have $x^2(1-x)(1-n) > 0$ and we can check linear independence by inspecting

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} g(x,n)(2-3x) - n(-1+(1+\kappa)x) - \left((1-2x)g(x,n) + x(1-x)\frac{\partial g}{\partial x}\right) \\ -n(1-n)(1+\kappa) \end{bmatrix}$$

Since the second entry of the right vector is always non-zero, these two vectors are linearly independent, and we conclude that u^* is a non-singular control. In other words, $\varphi(t) = \langle \boldsymbol{\lambda}, G \rangle$ and $\dot{\varphi}(t) = \langle \boldsymbol{\lambda}, [F,G] \rangle$ cannot both be zero at the same time. Therefore, there are no singular arcs for the controller u^* that satisfies the properties of Pontryagin's Maximum Principle.

4.6 Discussion

In this chapter, we extended the game-environment feedback model of [119] with a dynamic public opinion state. The public opinion imperfectly tracks the true environmental state, creating a lag between changes in the environment and changes in opinion. Due to this lag, the dynamics can become destabilized in regimes from the original model where an intermediate population and resource state is stable.

We leveraged this extension to investigate information-based policies that encourage the population to accumulate common resources. We considered two ways public opinion can be influenced. The first is a propaganda-like intervention where an external influencing agent attempts to sway public opinion. The second aims to accurately inform the public of the current true environmental state, e.g. through environmental

education programs or awareness campaigns. For each of these scenarios, we formulate an optimal control problem where the objective is to maximize accumulation of the environment state. We used a Hamiltonian-based optimal control algorithm [39] to numerically characterize locally optimal control strategies. We find that the propaganda strategies' ability to restore the environment in the cases of inevitable tragedy is limited. When control effort is costly, the return on investment from spreading information is not worth the benefit. In other dynamical regimes, these policies always bias public opinion towards lower environmental states to instigate cooperative behavior. However, because there are costs for control effort, the injection of information is done at strategic times. Similar conclusions hold for awareness-based policies, in which learning of the true environment is encouraged when public opinion overestimates the true state. We also demonstrated incentive-based policies induce a similar outcome. Hence, our simulation study points to a general scheme of influencing public opinion to accumulate common resources - drive the system to an oscillating tragedy of the commons by concentrating influence effort near the critical points of the dynamics.

The resulting oscillating tragedy of the commons maximizes accumulation of common resources over a finite time horizon because the policies increase the amount of time spent at replete states. The major drawback is the common resource will collapse. These effects are extremely undesirable if there are no alternative resource options. Hence, different ways of thinking about control are necessary. How should objective functions be specified to reduce extreme variation while maintaining stability of the common resource? Perhaps priority needs to be placed on the population's well-being in addition to conserving the commons [59]. The information policies we studied generally are unable to revive the commons in regimes where tragedy is inevitable. It remains to show whether combination policies such as influencing the population's opinion and incentive are sufficient to avert a tragedy of the commons.

Such policies are called for by researchers in the social sciences, who argue that information and incentives are each individually necessary, but insufficient [82, 113]

Our formulation of public opinion serves to describe how a population's overall perception of the environment can dynamically change over time. Here, the individual payoffs are modulated by perception and hence agents' decisions are solely driven by controlled beliefs. Specific aspects of an individual's decision-making process are abstracted here. We point the interested reader to relevant works that study these considerations in common-pool resource scenarios [60, 43], where separate game-theoretic and cognitive analyses are called for. In our model, the agents experience perceived payoffs. An alternate formulation may involve the partially informed agents changing their environmental opinions, and hence future action, upon experiencing *actual* (n-dependent) payoffs. This presents an opportunity for further study - opinionated agents with payoff-based reasoning capabilities can shape how perceptions change, and in turn, the application of information-based control.

Appendix

4.A Stability proofs of dynamical regimes

The proofs in this Appendix section are adapted from the Supplementary materials of [119].

Proof of Proposition 4.1 - periodic orbits

To prove this result, we need to consider the following energy function,

$$M(x,n) = -\delta_{PS}\left[\log n + \log(1-n)\right] - \log x - \frac{\theta}{1+\Delta}\log(1-x) + \frac{1+\Delta+\theta}{1+\Delta}\log(1+\Delta x)$$

where $\delta_{PS} = P_1 - S_1$, and $\delta_{TR} = T_1 - R_1$, and $\Delta = \frac{\delta_{TR}}{\delta_{PS}} - 1$. We state the following lemma, regarding the level curves of H.

Lemma 4.1. *For an arbitrary point $(x_0, n_0) \in (0,1)^2$, the set*

$$\{(x,n) \in (0,1)^2 : M(x,n) = M(x_0,n_0)\}$$

is given by a closed curve parameterized symmetrically about $n = 1/2$ by the top and bottom halves,

$$n_{\pm}(x) = \frac{1}{2}\left(1 \pm \sqrt{1 - 4\exp\left(\frac{M(x_0,n_0) - B(\theta,\Delta,x)}{\delta_{PS}}\right)}\right). \tag{4.13}$$

where

$$B(\theta,\Delta,x) = \log x + \frac{\theta}{1+\Delta}\log(1-x) - \frac{1+\Delta+\theta}{1+\Delta}\log(1+\Delta x)$$

Proof. The parameterization $n_{\pm}(x)$ is obtained by setting $M(x_0,n_0) = M(x,n)$, and solving for n. We get

$$\frac{M(x_0,n_0) - B(\theta,\Delta,x)}{\delta_{PS}} = \log n + \log(1-n)$$

Taking the exponential and using the quadratic equation, we obtain (4.13). ■

Proof of Proposition 4.1. Under the conditions of this result (4.5), the $\dot{x}$ equation can be written as

$$\dot{x} = \delta_{PS}(1-2n)x(1-x)(1+\Delta x)$$

For any $(x,n) \in (0,1)^2$, we have

$$\dot{M}(x,n) = \frac{\partial M}{\partial x}\dot{x} + \frac{\partial M}{\partial n}\dot{n} = 0$$

To see this, we compute

$$\frac{\partial M}{\partial x} = \frac{1}{x} - \frac{\theta}{1+\Delta}\frac{1}{1-x} - \frac{1+\Delta+\theta}{1+\Delta}\frac{\Delta}{1+\Delta x}$$
$$\frac{\partial M}{\partial n} = \delta_{PS}\left(\frac{1}{n} - \frac{1}{1-n}\right)$$
$$\frac{\partial M}{\partial x}\dot{x} = \delta_{PS}(1-2n)\left[(1-x)(1+\Delta x) - \frac{\theta}{1+\Delta}x(1+\Delta x) - \frac{1+\Delta+\theta}{1+\Delta}\Delta x(1-x)\right]$$

The x^2 term inside the square bracket is zero, the x^1 term is $-(1+\theta)$, and the x^0 term is 1. Hence, $\frac{\partial M}{\partial x}\dot{x} = -\frac{\partial M}{\partial n}\dot{n} = -\delta_{PS}(1-2n)(-1+(1+\theta)x)$. Therefore, $M(x(t),n(t)) = M(x_0,n_0)$ for all $t \geq 0$ and the trajectory $(x(t),n(t))$ follows the periodic orbit given in Lemma 4.1. ■

Proof of Proposition 4.2 - asymptotically stable heteroclinic cycle

Proof of Proposition 4.2. Under the condition (4.6), the four corner fixed points are saddles and the interior fixed point is unstable. The proof relies on properties of the characteristic matrix C of the heteroclinic cycle Λ (see Section 2, [42]), which encodes its stability properties and is constructed as follows.

Observe that $X = [0,1]^2$ is the intersection of four halfspaces,

$$X = \{x \geq 0\} \cap \{1-n \geq 0\} \cap \{n \geq 0\} \cap \{1-x \geq 0\}.$$

Define $x_1 = x$, $x_2 = 1-n$, $x_3 = n$, $x_4 = 1-x$. We can write the augmented dynamics

as (with some abuse of notation for the function g (4.3))

$$\dot{x_1} = x_1 x_4 g(x_1, x_2, x_3, x_4)$$
$$\dot{x_2} = -x_2 x_3(-1 + (1+\theta)x_1)$$
$$\dot{x_3} = x_2 x_3(-1 + (1+\theta)x_1)$$
$$\dot{x_4} = -x_1 x_4 g(x_1, x_2, x_3, x_4)$$

The characteristic matrix C has elements defined by

$$C_{ij} = \left.\frac{\dot{x}_j}{x_j}\right|_{z_i}, \; i, j = 1, 2, 3, 4$$

where z_i is one of the four corner fixed points. We obtain C by following the scheme below.

	x_1	x_2	x_3	x_4
(0,0)	$S_0 - P_0$	0	-1	0
(1,0)	$S_1 - P_1$	1	0	0
(0,1)	0	0	θ	$T_0 - R_0$
(1,1)	0	$-\theta$	0	$T_1 - R_1$

Under the condition of the Proposition, $S_0 - P_0 > 0$, $S_1 - P_1 < 0$, $T_0 - R_0 < 0$, and $T_1 - R_1 > 0$. Hence, C has positive entries occuring only on the diagonal, and each row and column contains only one positive entry. Thus, Λ is a *simple heteroclinic cycle* [42]. We review some technical conditions for the stability of heteroclinic cycles. From section 4 in [42], it is required that

- The heteroclinic cycle Λ is a compact invariant subset of ∂X, the boundary of X.
- Let $\Lambda_k \subset \Lambda$ for $k = 1, \ldots, m$ (m is the number of fixed points in Λ) be such that for each $x \in \Lambda$, there is a k with $\omega(x) \subset \Lambda_k$ ($\omega(x)$ is the ω-limit set of x).

The above conditions are satisfied in our case since $\Lambda = \partial X$ is compact, ∂X is invariant under the dynamics, and the four corner fixed points serve as the Λ_k's. The stability of Λ invokes the following result from [42]:

Lemma 4.2 (Corollary 1 in [42])**.** *Let Λ be a simple heteroclinic cycle, which is asymptotically stable within ∂X. (This is automatically satisfied if the cycle is robust and all Λ_k are fixed points.) Then C is a square matrix (after elimination of superfluous columns) with positive entries occuring only in the main diagonal (after a suitable rearrangement of the rows or columns). Let $\det C \neq 0$. If C is not an M-matrix (at least one leading principal minor is negative) then Λ is asymptotically stable.*

Indeed, C is not an M-matrix by the following calculation and invoking condition (4.6):

$$\det C = \theta \Big[(S_0 - P_0)(T_1 - R_1) - (P_1 - S_1)(R_0 - T_0) \Big] < 0$$

■

Three other dynamical regimes

Here, we will use the Jacobian DF of the system dynamics F, which is given by

$$DF(x,n) = \begin{bmatrix} (1-2x)g(x,n) + \frac{\partial g}{\partial x}(n)x(1-x) & x(1-x)\frac{\partial g}{\partial n}(x) \\ n(1-n)(1+\theta) & (1-2n)((1+\theta)x - 1) \end{bmatrix}$$

At the corner equilibria, we have

$$\begin{aligned} DF(0,0) &= \begin{bmatrix} S_0 - P_0 & 0 \\ 0 & -1 \end{bmatrix}, \quad DF(1,0) = \begin{bmatrix} T_0 - R_0 & 0 \\ 0 & \theta \end{bmatrix} \\ DF(0,1) &= \begin{bmatrix} S_1 - P_1 & 0 \\ 0 & 1 \end{bmatrix}, \quad DF(1,1) = \begin{bmatrix} T_1 - R_1 & 0 \\ 0 & -\theta \end{bmatrix} \end{aligned} \tag{4.14}$$

Under certain conditions, there is an interior fixed point given by

$$(x^*, n^*) = \left(\frac{1}{1+\theta}, \frac{\theta(P_0 - S_0) + (T_0 - R_0)}{\theta(P_0 - S_0) + (T_0 - R_0) + \theta(S_1 - P_1) + (R_1 - T_1)} \right) \tag{4.15}$$

We will consider three cases for the payoff structure of A_0: when $T_0 > R_0, S_0 > P_0$ and $T_0 < R_0, S_0 < P_0$.

Case 1: $R_0 - T_0 < 0, S_0 - P_0 > 0$ **(regions TOC4 and V1)**

In this payoff structure, the game A_0 has a mixed Nash at

$$x_m = \frac{S_0 - P_0}{(S_0 - P_0) + (T_0 - R_0)} \in (0, 1)$$

and From the Jacobian analysis, all corner points are unstable fixed points. The following result characterizes the stability properties of h.

Theorem 4.1. *If $x_m < \frac{1}{1+\theta}$, $(x_m, 0)$ is a stable fixed point (region TOC4). If $x_m > \frac{1}{1+\theta}$, $(x_m, 0)$ is unstable and (x^*, n^*) is stable (region V2).*

Proof. We have

$$DF(x, n) = \begin{bmatrix} \frac{(T_0 - R_0)(P_0 - S_0)}{(S_0 - P_0) + (T_0 - R_0)} & x_m(1 - x_m)\frac{\partial g}{\partial n}(x_m) \\ 0 & (1 + \theta)x_m - 1 \end{bmatrix} \tag{4.16}$$

The eigenvalues λ_1, λ_2 are the diagonal elements of the above. We already have $\lambda_1 < 0$ since $S_0 > P_0$. To ensure stability, we need $(1 + \theta)x_m - 1 < 0$, or when $x_m < \frac{1}{1+\theta}$.

For the second claim, we have $x_m > \frac{1}{1+\theta}$, or $\theta(P_0 - S_0) + (T_0 - R_0) < 0$. This ensures that $n^* \in (0, 1)$, since $\theta(S_1 - P_1) + (R_1 - T_1) < 0$. Also,

$$DF(x^*, n^*) = \begin{bmatrix} \frac{\partial g}{\partial x}(n^*)\frac{\theta}{(1+\theta)^2} & \frac{\partial g}{\partial n}(x^*)\frac{\theta}{(1+\theta)^2} \\ n^*(1 - n^*)(1 + \theta) & 0 \end{bmatrix} \tag{4.17}$$

The eigenvalues are

$$\frac{1}{2}\left[\frac{\partial g}{\partial x}(n^*)\frac{\theta}{(1+\theta)^2} \pm \sqrt{\left(\frac{\partial g}{\partial x}(n^*)\frac{\theta}{(1+\theta)^2}\right)^2 + 4\frac{\partial g}{\partial n}(x^*)\frac{\theta n^*(1 - n^*)}{1 + \theta}}\right] \tag{4.18}$$

The fixed point (x^*, n^*) is stable if both eigenvalues have negative real parts. This is the case if $\frac{dg}{dx}(n^*), \frac{dg}{dn}(x^*) < 0$. Indeed, from (4.3), we need

$$\frac{\partial g}{\partial x}(n^*) = (1 + \theta)\Big(n^*((P_1 - S_1) + (S_0 - P_0)) + (P_0 - S_0)\Big) < 0$$

or

$$\frac{\theta(S_0 - P_0)}{\theta(S_0 - P_0) + \theta(P_1 - S_1)} > \frac{\theta(S_0 - P_0) - (T_0 - R_0)}{\theta(S_0 - P_0) + \theta(P_1 - S_1) - (T_0 - R_0) + (T_1 - R_1)}$$

By inspection, the above is satisfied and therefore $\frac{\partial g}{\partial x}(n^*) < 0$. We also need

$$\frac{\partial g}{\partial n}(x^*) = \frac{1}{n^*}\Big((P_0 - S_0) + x^*((T_0 - R_0) + (S_0 - P_0))\Big) < 0$$

which reduces to $x_m > \frac{1}{1+\theta}$.

■

Case 2: $R_0 - T_0 > 0$**,** $S_0 - P_0 < 0$ **(regions TOC1 and TOC2)**

In this payoff structure, the game A_0 has a mixed Nash at

$$x_m = \frac{P_0 - S_0}{(P_0 - S_0) + (R_0 - T_0)} \in (0, 1)$$

From the Jacobian analysis in (4.14), the corner point $(0, 0)$ is stable and remaining corner points are unstable. The following result characterizes the stability properties of the other (mixed) equilibria of h.

Theorem 4.2. $(x_m, 0)$ *and* (x^*, n^*) *are unstable fixed points.*

Proof. At $(x_m, 0)$ the Jacobian have the same structure in (4.16). The first diagonal element of $DF(x, n)$ in (4.16) is positive in this case, implying $\lambda_1 > 0$. Hence, $(x_m, 0)$ is unstable.

Second, we consider (x^*, n^*) with the Jacobian in (4.17). First note that if $x_m > \frac{1}{1+\theta}$ then $n^* \notin (0, 1)$ by (4.15).

Next, we focus on the case $x_m < \frac{1}{1+\theta}$, which implies $\theta(P_0 - S_0) < R_0 - T_0$. In this case, the numerator of n^* is negative. The denominator of n^* is a smaller negative value because $\theta(S_1 - P_1) + R_1 - T_1 < 0$ hence $n^* \in (0, 1)$. We consider the condition $\frac{dg}{dx}(n^*) > 0$, which is given by

$$n^*((P_1 - S_1) + (S_0 - P_0)) > S_0 - P_0, \tag{4.19}$$

to show eigenvalues (4.18) are positive. Note here that if $P_0 - S_0 < P_1 - S_1$ then the left hand side is positive because $n^* > 0$. Hence the condition in (4.19) is true. When the reverse holds, that is, $P_0 - S_0 > P_1 - S_1$, then the condition above becomes

$$n^* < \frac{(S_0 - P_0)}{(P_1 - S_1) + (S_0 - P_0)} = \frac{\theta(P_0 - S_0)}{\theta(S_1 - P_1) + \theta(P_0 - S_0)}$$

Note here that the right hand side is greater than 1 because $S_1 - P_1 < 0$. Hence the condition above holds. Therefore, $\frac{\partial g}{\partial x}(n^*) > 0$ when $x_m < \frac{1}{1+\theta}$. ■

CHAPTER V

CONCLUSIONS AND FUTURE WORK

This thesis studied the role of information in three distinct multi-agent systems that appear in social and biological contexts: 1) collective behaviors, 2) epidemic awareness, and 3) averting tragedies of the commons. In these systems, studying how information is distributed among the population is essential to understanding why global behaviors emerge. Our work furthers this understanding in their respective application domains, and raises important questions for influence and control. We outline these contributions and insights for the three systems below.

The role of information in collective behavior

Collective behaviors are central aspects of evolutionary biology, microbial decision making, and animal communication [120, 22]. We showed how information sharing promotes coordination in collective decision-making. In particular, when the quality of communication is high, individuals maximize chances to coordinate in fluctuating environments when they adopt "majority logic" strategies that base decisions on shared information. Because these strategies were identified from a first principles approach, they offer insight into the mechanisms that maintain coordinated group behaviors under the limitations of a noisy communication system. These insights contribute to advancing core principles of how collective decisions are made.

A potential future application of these principles is the development of alternative antibiotic therapies. In particular, biofilm formation may be inhibited by disrupting quorum sensing communication systems in bacteria. Biofilms are extra-cellular

bacterial structures that have decreased susceptibility to antibiotic drugs and immune responses. In the context of our work, inhibition can conceptually be achieved by degrading the quality of shared signals, i.e., autoinducer signalling molecules. A promising approach is to develop synthetic enzymes that break down autoinducers [71]. However, many questions remain. What stage of infection should such therapies be implemented? Which pathogenic bacteria can be targeted? Significant further efforts to develop theoretical and empirical understanding of these complex interactions are necessary before such therapies can be realized [96, 121].

Epidemic awareness

Informed individuals change their social behaviors to prevent getting sick from infectious diseases [35, 14]. We analyzed to what extent a linear form of individual awareness mitigates epidemic spreading. Our results point to the effectiveness as well as limitations of social distancing measures. Specifically, we found that in our formulation, awareness-induced social distancing alone is insufficient to eradicate an epidemic, though it lowers the endemic equilibrium state. We also identified local contact information as the most effective source in slowing the spread of an epidemic, but not in eradicating the disease. These insights provide a deeper theoretical understanding of the outcomes of awareness-coupled disease epidemic dynamics, which can potentially help inform improved epidemic control policies [35, 50]. For example, disease forecasts can be refined with the help of more realistic model predictions [4], providing improved guidelines for resource employment for epidemic control. One epidemic intervention control measure that gathers local contact information is known as contact tracing, which involves informing and monitoring the people an infected individual comes into contact with. This process can be quite challenging, requiring costly investment on the part of centralized health organizations [6, 18]. At the start of an outbreak, it is critical to prevent the disease from spreading in the first place.

This requires a higher order level of awareness and social distancing not captured in our model. This provides motivation to improve the efficiency of contact tracing methods and the effectiveness of disease awareness campaigns.

Averting a tragedy of the commons

To sustain common resources, effective strategies to avert a tragedy of the commons are paramount [73, 113]. We showed how a dynamic public opinion can be influenced to promote the conservation of common resources. In particular, using an optimal control algorithm to numerically compute locally optimal controls, we found that these information policies will drive the resource to cycle between replete and deplete states in order to maximize accumulation of resources over time. These high variance effects are extremely undesirable because resource abundance is eventually followed by a collapse. Instead of finding policies that seek to maximize only resource abundance over time, a different objective should be specified that promotes *stability* of the commons. It remains to identify what types of policies would achieve this objective in our dynamical formulation. There is general agreement that utilizing a combination of both public information and incentive influence policies may be required to avert a tragedy of the commons [82]. Additionally, the distinction between perceived (opinion-dependent) incentives that determine actions and experienced (environment-dependent) payoffs that determine benefits is not captured in our model. A challenge for further study is to consider how these experienced payoffs may affect changes in environmental opinion, and in turn, how opinion control may be implemented.

Overall, this thesis provides new insights and principled approaches to predict the dynamical outcomes of complex, multi-agent systems. In doing so, this thesis has highlighted the role of information, awareness, and feedback, as a means to better characterize socio-biological multi-agent systems.

REFERENCES

[1] Ahn, H. J. and Hassibi, B., "Global dynamics of epidemic spread over complex networks," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pp. 4579–4585, Dec 2013.

[2] Ahn, H. J. and Hassibi, B., "Global dynamics of epidemic spread over complex networks," in *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on*, pp. 4579–4585, Dec 2013.

[3] Ahn, H. J. and Hassibi, B., "On the mixing time of the sis markov chain model for epidemic spread," in *Decision and Control (CDC), 2014 IEEE 53rd Annual Conference on*, pp. 6221–6227, Dec 2014.

[4] Anderson, R. M., May, R. M., and Anderson, B., *Infectious diseases of humans: dynamics and control*, vol. 28. Oxford University Press, 1992.

[5] Arganda, S., Pérez-Escudero, A., and de Polavieja, G. G., "A common rule for decision making in animal collectives across species," *Proceedings of the National Academy of Sciences*, vol. 109, no. 50, pp. 20508–20513, 2012.

[6] Armbruster, B. and Brandeau, M. L., "Contact tracing to control infectious disease: when enough is enough," *Health Care Management Science*, vol. 10, pp. 341–355, Dec 2007.

[7] Armijo, L., "Minimization of functions having lipschitz continuous first partial derivatives.," *Pacific J. Math.*, vol. 16, no. 1, pp. 1–3, 1966.

[8] Aronson, J., Milton, S., and Blignaut, J., *Restoring Natural Capital: Science, Business, and Practice*. Island Press, 2007.

[9] Atkinson, S., Chang, C.-Y., Sockett, R. E., Cámara, M., and Williams, P., "Quorum sensing in yersinia enterocolitica controls swimming and swarming motility," *Journal of Bacteriology*, vol. 188, no. 4, pp. 1451–1461, 2006.

[10] Axelrod, R., *The Evolution of Cooperation*. Basic Books, 1984.

[11] Backwell, P. R. Y., Christy, J. H., Telford, S. R., Jennions, M. D., and Passmore, N. I., "Dishonest signalling in a fiddler crab," *Proceedings: Biological Sciences*, vol. 267, no. 1444, pp. 719–724, 2000.

[12] Barabási, A.-L. and Albert, R., "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[13] Basar, T. and Olsder, G. J., *Dynamic noncooperative game theory*. Siam, 1999.

[14] Bauch, C. T. and Galvani, A. P., "Social factors in epidemiology," *Science*, vol. 342, no. 6154, pp. 47–49, 2013.

[15] Bergstrom, C. T. and Lachmann, M., "Signalling among relatives. i. is costly signalling too costly?," *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 352, no. 1353, pp. 609–617, 1997.

[16] Brown, S. P. and Johnstone, R. A., "Cooperation in the dark: Signalling and collective action in quorum-sensing bacteria," *Proceedings of the Royal Society B: Biological Sciences*, vol. 268, no. 1470, pp. 961–965, 2001.

[17] Brown, S. P. and Taddei, F., "The durability of public goods changes the dynamics and nature of social dilemmas," *PLOS ONE*, vol. 2, pp. 1–7, 07 2007.

[18] Browne, C., Gulbudak, H., and Webb, G., "Modeling contact tracing in outbreaks with application to ebola," *Journal of Theoretical Biology*, vol. 384, pp. 33 – 49, 2015.

[19] Burgos, A. C. and Polani, D., "Cooperation and antagonism in information exchange in a growth scenario with two species," *Journal of Theoretical Biology*, vol. 399, pp. 117 – 133, 2016.

[20] Cohen, J., "Population growth and earth's human carrying capacity," *Science*, vol. 269, no. 5222, pp. 341–346, 1995.

[21] Cornforth, D. M., Reluga, T. C., Shim, E., Bauch, C. T., Galvani, A. P., and Meyers, L. A., "Erratic flu vaccination emerges from short-sighted behavior in contact networks," *PLOS Computational Biology*, vol. 7, pp. 1–10, 01 2011.

[22] Couzin, I. D., "Collective cognition in animal groups," *Trends in Cognitive Sciences*, vol. 13, no. 1, pp. 36 – 43, 2009.

[23] Couzin, I. D., Ioannou, C. C., Demirel, G., Gross, T., Torney, C. J., Hartnett, A., Conradt, L., Levin, S. A., and Leonard, N. E., "Uninformed individuals promote democratic consensus in animal groups," *Science*, vol. 334, no. 6062, pp. 1578–1580, 2011.

[24] Damore, J. A. and Gore, J., "Understanding microbial cooperation," *Journal of Theoretical Biology*, vol. 299, pp. 31 – 41, 2012.

[25] Dawes, R. M., C, I. I., Hi, H. I., and Dawes, R. M., "Formal models of dilemmas in social decision-making," in *H u m n Judgment and Decision Processes*, Amdemic Press, 1975.

[26] de Kievit, T. R. and Iglewski, B. H., "Bacterial quorum sensing in pathogenic relationships," *Infection and Immunity*, vol. 68, no. 9, pp. 4839–4849, 2000.

[27] Donaldson-Matasci, M. C., Bergstrom, C. T., and Lachmann, M., "The fitness value of information," *Oikos*, vol. 119, no. 2, pp. 219–230, 2010.

[28] Donaldson-Matasci, M. C., Bergstrom, C. T., and Lachmann, M., "When unreliable cues are good enough," *The American Naturalist*, vol. 182, pp. 313–327, Sept 2013.

[29] Donaldson-Matasci, M. C., Lachmann, M., and Bergstrom, C. T., "Phenotypic diversity as an adaptation to environmental uncertainty," *Evolutionary Ecology Research*, vol. 10, pp. 493–515, 2008.

[30] Eguíluz, V. M. and Klemm, K., "Epidemic threshold in structured scale-free networks," *Phys. Rev. Lett.*, vol. 89, p. 108701, Aug 2002.

[31] Eksin, C., Shamma, J. S., and Weitz, J. S., "Disease dynamics in a stochastic network game: a little empathy goes a long way in averting outbreaks," vol. 7, pp. 44122 EP –, 03 2017.

[32] Franks, N. R., Pratt, S. C., Mallon, E. B., Britton, N. F., and Sumpter, D. J. T., "Information flow, opinion polling and collective intelligence in house–hunting social insects," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 357, no. 1427, pp. 1567–1583, 2002.

[33] Funk, S., Bansal, S., Bauch, C. T., Eames, K. T., Edmunds, W. J., Galvani, A. P., and Klepac, P., "Nine challenges in incorporating the dynamics of behaviour in infectious diseases models," *Epidemics*, vol. 10, no. Supplement C, pp. 21 – 25, 2015. Challenges in Modelling Infectious DIsease Dynamics.

[34] Funk, S., Gilad, E., Watkins, C., and Jansen, V. A. A., "The spread of awareness and its impact on epidemic outbreaks," *Proceedings of The National Academy of Sciences*, vol. 106, pp. 6872–6877, 2009.

[35] Funk, S., Salathé, M., and Jansen, V. A. A., "Modelling the influence of human behaviour on the spread of infectious diseases: a review," *Journal of The Royal Society Interface*, 2010.

[36] Ganesh, A., Massoulie, L., and Towsley, D., "The effect of network topology on the spread of epidemics," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 2, pp. 1455–1466 vol. 2, March 2005.

[37] Granell, C., Gómez, S., and Arenas, A., "Dynamical interplay between awareness and epidemic spreading in multiplex networks," *Phys. Rev. Lett.*, vol. 111, p. 128701, Sep 2013.

[38] Granell, C., Gómez, S., and Arenas, A., "Competing spreading processes on multiplex networks: Awareness and epidemics," *Phys. Rev. E*, vol. 90, p. 012808, Jul 2014.

[39] Hale, M., Wardi, Y., Jaleel, H., and Egerstedt, M., "Hamiltonian-based algorithm for optimal control," *Submitted for publication*, 2016.

[40] Hardin, G., "The tragedy of the commons," *Science*, vol. 162, no. 3859, pp. 1243–1248, 1968.

[41] Henke, J. M. and Bassler, B. L., "Bacterial social engagements," *Trends in Cell Biology*, vol. 14, no. 11, pp. 648–656, 2004.

[42] Hofbauer, J., "Heteroclinic cycles in ecological differential equations," in *Equadiff 8*, pp. 105–116, Mathematical Institute, Slovak Academy of Sciences, 1994.

[43] Hota, A. R., Garg, S., and Sundaram, S., "Fragility of the commons under prospect-theoretic risk attitudes," *Games and Economic Behavior*, vol. 98, pp. 135 – 164, 2016.

[44] Hota, A. R. and Sundaram, S., "Controlling human utilization of shared resources via taxes," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 6984–6989, Dec 2016.

[45] Hota, A. R. and Sundaram, S., "Game-theoretic protection against networked SIS epidemics by human decision-makers," *CoRR*, vol. abs/1703.08750, 2017.

[46] Hutchings, J. A. and Reynolds, J. D., "Marine fish population collapses: Consequences for recovery and extinction risk," *BioScience*, vol. 54, no. 4, pp. 297–309, 2004.

[47] Huttegger, S., Skyrms, B., Tarrés, P., and Wagner, E., "Some dynamics of signaling games," *Proceedings of the National Academy of Sciences*, vol. 111, no. Supplement 3, pp. 10873–10880, 2014.

[48] Istratescu, V. I., *Fixed Point Theory, An Introduction*. Holland: D.Reidel, 1981.

[49] Jablonka, E., "Information: Its interpretation, its inheritance, and its sharing," *Philosophy of Science*, vol. 69, no. 4, pp. 578–605, 2002.

[50] Keeling, M. and Rohani, P., *Modeling Infectious Diseases in Humans and Animals*. Princeton, NJ: Princeton University Press, 2011.

[51] Kermack, W. and McKendrick, A., "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 115, no. 772, pp. 700–721, 1927.

[52] King, A. J. and Cowlishaw, G., "When to use social information: the advantage of large group size in individual decision making," *Biology Letters*, vol. 3, no. 2, pp. 137–139, 2007.

[53] Kussell, E. and Leibler, S., "Phenotypic diversity, population growth, and information in fluctuating environments," *Science*, vol. 309, no. 5743, pp. 2075–2078, 2005.

[54] Lachmann, M., Sella, G., and Jablonka, E., "On the advantages of information sharing," *Proceedings of the Royal Society of London B: Biological Sciences*, vol. 267, no. 1450, pp. 1287–1293, 2000.

[55] Liberzon, D., *Calculus of Variations and Optimal Control Theory: A Concise Introduction*. Princeton University Press, 2012.

[56] Lindvall, T., *Lectures on the Coupling Method*. Dover Books on Mathematics Series, Dover Publications, Incorporated, 2002.

[57] Lopez, D., Vlamakis, H., and Kolter, R., "Generation of multiple cell types in bacillus subtilis," *FEMS Microbiology Reviews*, vol. 33, no. 1, pp. 152–163, 2009.

[58] Lupia, A., "Communicating science in politicized environments," *Proceedings of the National Academy of Sciences*, vol. 110, no. Supplement 3, pp. 14048–14054, 2013.

[59] Manzoor, T., Aseev, S., Rovenskaya, E., and Muhammad, A., "Optimal control for sustainable consumption of natural resources," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 10725 – 10730, 2014. 19th IFAC World Congress.

[60] Manzoor, T., E., R., and Muhammad, A., "Game-theoretic insights into the role of environmentalism and social-ecological relevance: A cognitive model of resource consumption," *Ecological Modelling*, vol. 340, pp. 74 – 85, 2016.

[61] Meacham, F., Perlmutter, A., and Bergstrom, C. T., "Honest signalling with costly gambles," *Journal of The Royal Society Interface*, vol. 10, no. 87, 2013.

[62] Mieghem, P. V., Omic, J., and Kooij, R., "Virus spread in networks," *IEEE/ACM Transactions on Networking*, vol. 17, pp. 1–14, Feb 2009.

[63] Miller, M. B., , and Bassler, B. L., "Quorum sensing in bacteria," *Annual Review of Microbiology*, vol. 55, no. 1, pp. 165–199, 2001. PMID: 11544353.

[64] Miller, N., Garnier, S., Hartnett, A. T., and Couzin, I. D., "Both information and social cohesion determine collective decisions in animal groups," *Proceedings of the National Academy of Sciences*, vol. 110, no. 13, pp. 5263–5268, 2013.

[65] Molina, C. and Earn, D. J. D., "Game theory of pre-emptive vaccination before bioterrorism or accidental release of smallpox," *Journal of The Royal Society Interface*, vol. 12, no. 107, 2015.

[66] Monderer, D. and Shapley, L. S., "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124 – 143, 1996.

[67] Nadell, C. D., Xavier, J. B., Levin, S. A., and Foster, K. R., "The evolution of quorum sensing in bacterial biofilms," *PLOS Biology*, vol. 6, pp. 1–9, 01 2008.

[68] Ndeffo Mbah, M. L., Liu, J., Bauch, C. T., Tekel, Y. I., Medlock, J., Meyers, L. A., and Galvani, A. P., "The impact of imitation on vaccination behavior in social contact networks," *PLoS Comput Biol*, vol. 8, p. e1002469, 04 2012.

[69] Newman, M., *Networks: An Introduction*. New York, NY, USA: Oxford University Press, Inc., 2010.

[70] Ogura, M. and Preciado, V. M., "Epidemic processes over adaptive state-dependent networks," *Phys. Rev. E*, vol. 93, p. 062316, Jun 2016.

[71] O'Loughlin, C. T., Miller, L. C., Siryaporn, A., Drescher, K., Semmelhack, M. F., and Bassler, B. L., "A quorum-sensing inhibitor blocks pseudomonas aeruginosa virulence and biofilm formation," *Proceedings of the National Academy of Sciences*, vol. 110, p. 17981, 10 2013.

[72] Omic, J., *Epidemics in networks: modeling, optimization and security games*. PhD thesis, Delft University of Technology, 2010.

[73] Ostrom, E., *Governing the Commons: The evolution of institutions for collective action*. Cambridge University Press, 1990.

[74] Paarporn, K., Eksin, C., and Weitz, J. S., "Information sharing for a coordination game in fluctuating environments," *bioRxiv*, 2018.

[75] Paarporn, K., Eksin, C., Weitz, J. S., and Shamma, J. S., "Epidemic spread over networks with agent awareness and social distancing," in *2015*

53rd Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 51–57, Sept 2015.

[76] Paarporn, K., Eksin, C., Weitz, J. S., and Shamma, J. S., "The effect of awareness on networked sis epidemics," in *Decision and Control (CDC), 2016 IEEE 55th Annual Conference on (submitted)*, Dec 2016.

[77] Paarporn, K., Eksin, C., Weitz, J. S., and Shamma, J. S., "Networked sis epidemics with awareness," *IEEE Transactions on Computational Social Systems*, vol. 4, pp. 93–103, Sept 2017.

[78] Paarporn, K., Eksin, C., Weitz, J. S., and Wardi, Y., "Optimal control policies for evolutionary dynamics with environmental feedback," *ArXiv*, March 2018.

[79] Pacheco, J. M., Vasconcelos, V. V., Santos, F. C., and Skyrms, B., "Co-evolutionary dynamics of collective action with signaling for a quorum," *PLOS Computational Biology*, vol. 11, pp. 1–12, 02 2015.

[80] Pandey, A., Atkins, K. E., Medlock, J., Wenzel, N., Townsend, J. P., Childs, J. E., Nyenswah, T. G., Ndeffo-Mbah, M. L., and Galvani, A. P., "Strategies for containing ebola in west africa," *Science*, vol. 346, no. 6212, pp. 991–995, 2014.

[81] Pastor-Satorras, R. and Vespignani, A., "Epidemic spreading in scale-free networks," *Phys. Rev. Lett.*, vol. 86, pp. 3200–3203, Apr 2001.

[82] Penn, D., "The evolutionary roots of our environmental problems: Toward a darwinian ecology," *The Quarterly Review of Biology*, vol. 78, no. 3, pp. 275–301, 2003.

[83] Pérez-Escudero, A. and de Polavieja, G. G., "Collective animal behavior from bayesian estimation and probability matching," *PLOS Computational Biology*, vol. 7, pp. 1–14, 11 2011.

[84] Perkins, T. J. and Swain, P. S., "Strategies for cellular decision-making," *Molecular Systems Biology*, vol. 5, no. 1, 2009.

[85] Perra, N., Balcan, D., Gonçalves, B., and Vespignani, A., "Towards a characterization of behavior-disease models," *PLoS ONE*, vol. 6, p. e23084, 08 2011.

[86] Popat, R., Cornforth, D. M., McNally, L., and Brown, S. P., "Collective sensing and collective responses in quorum-sensing bacteria," *Journal of The Royal Society Interface*, vol. 12, no. 103, 2015.

[87] Pratt, S. C., Mallon, E. B., Sumpter, D. J., and Franks, N. R., "Quorum sensing, recruitment, and collective decision-making during colony emigration by the ant leptothorax albipennis," *Behavioral Ecology and Sociobiology*, vol. 52, pp. 117–127, Jul 2002.

[88] Press, W. H. and Dyson, F. J., "Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent," *Proceedings of the National Academy of Sciences*, vol. 109, no. 26, pp. 10409–10413, 2012.

[89] Rankin, D., Bargum, K., and Kokko, H., "The tragedy of the commons in evolutionary biology," *Trends in Ecology and Evolution*, vol. 22, no. 12, pp. 643 – 651, 2007.

[90] Rankin, D. and López-Sepulcre, A., "Can adaptation lead to extinction?," *Oikos*, vol. 111, no. 3, pp. 616–619, 2005.

[91] Rapoport, A. and Suleiman, R., "Equilibrium solutions for resource dilemmas," *Group Decision and Negotiation*, vol. 1, pp. 269–294, Nov 1992.

[92] Reluga, T. C., "Game theory of social distancing in response to an epidemic," *PLoS Comput Biol*, vol. 6, p. e1000793, 05 2010.

[93] Rivoire, O. and Leibler, S., "The value of information for populations in varying environments," *Journal of Statistical Physics*, vol. 142, no. 6, pp. 1124–1166, 2011.

[94] Rizzo, A., Frasca, M., and Porfiri, M., "Effect of individual behavior on epidemic spreading in activity-driven networks," *Phys. Rev. E*, vol. 90, p. 042801, Oct 2014.

[95] Roopnarine, P., "Ecology and the tragedy of the commons," *Sustainability*, vol. 5, no. 2, pp. 749–773, 2013.

[96] Rutherford, S. T. and Bassler, B. L., "Bacterial quorum sensing: Its role in virulence and possibilities for its control," *Cold Spring Harbor Perspectives in Medicine*, vol. 2, p. a012427, 11 2012.

[97] Sahneh, F. D. and Scoglio, C., "Epidemic spread in human networks," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 3008–3013, Dec 2011.

[98] Sandholm, W. H., "Potential games with continuous player sets," *Journal of Economic Theory*, vol. 97, no. 1, pp. 81 – 108, 2001.

[99] Sandholm, W. H., *Population Games and Evolutionary Dynamics*. MIT Press, 2010.

[100] Sasaki, T. and Pratt, S. C., "The psychology of superorganisms: Collective decision making by insect societies," *Annual Review of Entomology*, vol. 63, no. 1, pp. 259–275, 2018.

[101] Shang, Y., "Discrete-time epidemic dynamics with awareness in random networks," *International Journal of Biomathematics*, vol. 06, no. 02, p. 1350007, 2013.

[102] Shang, Y., "Modeling epidemic spread with awareness and heterogeneous transmission rates in networks," *J Biol Phys*, vol. 39, no. 3, pp. 489–500, 2013.

[103] Shang, Y., "Degree distribution dynamics for disease spreading with individual awareness," *Journal of Systems Science and Complexity*, vol. 28, no. 1, pp. 96–104, 2015.

[104] Skyrms, B., *Signals: Evolution, Learning, and Information*. Oxford University Press, 2010.

[105] SteelFisher, G. K., Blendon, R. J., Bekheit, M. M., and Lubell, K., "The public's response to the 2009 h1n1 influenza pandemic," *New England Journal of Medicine*, vol. 362, no. 22, p. e65, 2010. PMID: 20484390.

[106] Stern, P. C., "New environmental theories: Toward a coherent theory of environmentally significant behavior," *Journal of Social Issues*, vol. 56, no. 3, pp. 407–424, 2000.

[107] Stewart, A. J. and Plotkin, J. B., "From extortion to generosity, evolution in the iterated prisoner's dilemma," *Proceedings of the National Academy of Sciences*, vol. 110, no. 38, pp. 15348–15353, 2013.

[108] van Bergen, Y., Coolen, I., and Laland, K. N., "Nine-spined sticklebacks exploit the most reliable source when public and private information

conflict.," *Proceedings of the Royal Society B: Biological Sciences*, vol. 271, no. 1542, pp. 957–962, 2004.

[109] Van Mieghem, P., Omic, J., and Kooij, R., "Virus spread in networks," *Networking, IEEE/ACM Transactions on*, vol. 17, pp. 1–14, Feb 2009.

[110] van Vugt, M. and Samuelson, C. D., "The impact of personal metering in the management of a natural resource crisis: A social dilemma analysis," *Personality and Social Psychology Bulletin*, vol. 25, no. 6, pp. 735–750, 1999.

[111] Vega-Redondo, F., *Economics and the Theory of Games*. Cambridge University Press, 2003.

[112] Volz, E. M., Miller, J. C., Galvani, A., and Ancel Meyers, L., "Effects of heterogeneous and clustered contact patterns on infectious disease dynamics," *PLoS Comput Biol*, vol. 7, p. e1002042, 06 2011.

[113] Vugt, M. V., "Averting the tragedy of the commons: Using social psychological science to protect the environment," *Current Directions in Psychological Science*, vol. 18, no. 3, pp. 169–173, 2009.

[114] Wang, Y., Chakrabarti, D., Wang, C., and Faloutsos, C., "Epidemic spreading in real networks: an eigenvalue viewpoint," in *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.*, pp. 25–34, Oct 2003.

[115] Wang, Y., Chakrabarti, D., Wang, C., and Faloutsos, C., "Epidemic spreading in real networks: an eigenvalue viewpoint," in *Reliable Distributed Systems, 2003. Proceedings. 22nd International Symposium on*, pp. 25–34, Oct 2003.

[116] Wang, Z., Andrews, M. A., Wu, Z.-X., Wang, L., and Bauch, C. T., "Coupled disease–behavior dynamics on complex networks: A review," *Physics of Life Reviews*, vol. 15, pp. 1 – 29, 2015.

[117] Watts, D. J. and Strogatz, S. H., "Collective dynamics of'small-world'networks.," *Nature*, vol. 393, no. 6684, pp. 409–10, 1998.

[118] Weibull, J., *Evolutionary Game Theory*. MIT Press, 1997.

[119] Weitz, J. S., Eksin, C., Paarporn, K., Brown, S. P., and Ratcliff, W. C., "An oscillating tragedy of the commons in replicator dynamics with game-environment feedback," *Proceedings of the National Academy of Sciences*, vol. 113, no. 47, pp. E7518–E7525, 2016.

[120] West, S. A., Diggle, S. P., Buckling, A., Gardner, A., and Griffin, A. S., "The social lives of microbes," *Annual Review of Ecology, Evolution, and Systematics*, vol. 38, no. 1, pp. 53–77, 2007.

[121] Whiteley, M., Diggle, S. P., and Greenberg, E. P., "Progress in and promise of bacterial quorum sensing research," *Nature*, vol. 551, pp. 313–320, 2017.

[122] Wu, Q., Fu, X., Small, M., and Xu, X.-J., "The impact of awareness on epidemic spreading in networks," *Chaos*, vol. 22, no. 1, 2012.

[123] Zhang, H.-F., Xie, J.-R., Tang, M., and Lai, Y.-C., "Suppression of epidemic spreading in complex networks by local information based behavioral responses," *Chaos*, vol. 24, no. 4, 2014.